PHONEOTYPIC MODELING OF HUMAN BEHAVIORS AND PROPENSITIES

By

GHASSAN F. BATI

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Professor Vivek K. Singh

And approved by

________________________________________

________________________________________

________________________________________

________________________________________

New Brunswick, New Jersey

January, 2019

## ABSTRACT OF THE DISSERTATION

PHONEOTYPIC MODELING OF HUMAN BEHAVIORS AND PROPENSITIES

by GHASSAN F. BATI

Dissertation Director:

Professor Vivek K. Singh

With the growth in mobile social networks, social internet of things, and cyber-physical-social systems, there is an ever growing need to model and understand human beings as they interact with other humans and socio-technical ecosystems. In this dissertation, we focus on modeling three core human concepts – trust propensity, altruism propensity, and interpersonal trust using mobile phone metadata. Traditional methods for understanding an individual's propensities and behaviors have been surveys and lab experiments. However, the growth of "personal big data", which includes the use of various personal ubiquitous devices, is allowing for human behaviors and propensities to be modeled via lower-cost, quick, automated methods. This dissertation proposes a new methodology to model human behaviors and propensities based on *phoneotypes* (phone-based observations of a combination of people's traits) that aims to complement traditional methods like surveys with a ubiquitous data-driven automated method. The analysis and modeling employ multiple deep and shallow machine learning algorithms and are based on two datasets - Rutgers Well-being Study and MIT friends and family dataset. Overall, the findings suggest that: (1) many phone-based features are associated with participant's altruism, trust, and interpersonal trust scores;

(2) phone-based prediction models for altruism, trust propensity, and interpersonal trust performed statistically significantly better than comparable demography-based models. This dissertation paves way to study the associations between human behavioral propensities and long-term “in the wild” socio-mobile behavior, and to utilize “personal big data” with shallow and deep machine learning approaches to model altruism, trust, and interpersonal trust. A better modeling approach for human beings will have multiple applications in fields like healthcare, well-being, and urban planning.

## Acknowledgments

“Then, to God belongs ⌜all⌝ praise - Lord of the heavens and Lord of the earth, Lord of the worlds. And to Him belongs ⌜all⌝ Majesty in the heavens and the earth, and He is the Exalted in Might, the Wise.” [Quran, 45:36&37].

“Whoever is not grateful to people, he is not grateful to God.” [Prophet Muhammad Peace and Blessings of God be upon Him].

I am certain that words are sometimes not enough to express and show gratitude, yet I will try my best.

First and foremost, I would like to take this opportunity to thank from the bottom of my heart all the very special individuals who have tremendously impacted my life so far – as a son, a sibling, a husband, a grandson, a nephew, a cousin, a relative, a friend, a student, and an instructor. It is these great people who have shaped who I am today and contributed directly or indirectly to this dissertation.

Next, very special thanks go to my great advisor Professor Vivek K. Singh who played an important role in guiding and advising me throughout the PhD research endeavor. He demonstrated to me how to be a great scholar and humble person simultaneously.

Also, I would like to thank my respected dissertation committee members, Professor Zoran Gajic (the best graduate director I have ever met), Professor Ivan Marsic, and Professor Pradeep Atrey for their help and support.

Last but surely not the least, I owe Umm Al-Qura University, Makkah, Saudi Arabia, a great debt of gratitude for providing the generous fellowship to pursue my graduate studies in the US since July 2009 to November 2018. Also, I would like to thank my colleagues at the behavioral informatics lab for assisting in conducting the Rutgers Well-being Study, processing its data, and the great discussions we have had.

## Dedication

"My Lord! Inspire me to ⌜always⌝ be thankful for Your favors which You blessed me and my parents with, and to do good deeds that please You. And instill righteousness in my offspring. I truly repent to You, and I truly submit ⌜to Your Will⌝." [Quran, 46:15].

This dissertation is dedicated to:

"My Dear Parents and My Beloved Wife".

May God gather us in this world as well as in His paradise in the company of the prophets, the people of truth, the martyrs, and the righteous.

Ameen

## Table of Contents

## List of Tables

## List of Figures

# Chapter 1

# Introduction

## 1.1. Motivations

With the growth in personal, mobile, and ubiquitous computing, increasingly larger aspects of human life are mediated by devices. Consequently, the data captured by such devices, i.e. the "personal big data", is enabling a datafication of the human life and facilitating the creation of a rich composite personas of different users [1]. With 1.4 billion smartphones and millions of quantified-self device users, more and more users keep track of their behavior, which has already shown value in the fields of healthcare, well-being, and urban planning [2]–[5].

The growth of such "personal big data" is allowing for human behaviors and propensities to be modeled via lower-cost, quick, automated methods. In fact, some researchers consider smartphones to be a "vast psychological questionnaire that we are constantly filling out, both consciously and unconsciously" [6] and a recent *smartphone psychology manifesto* states that "… smartphones could transform psychology even more profoundly than PCs and brain imaging did" [7, p. 1].

At the same time, with the growth in mobile social networks, social internet of things, and cyber-physical-social systems, there is an ever growing need to model and understand human beings as they interact with other humans and socio-technical ecosystems. In this dissertation, we focus on modeling three such concepts – trust, altruism, and interpersonal trust using "personal big data".

Altruism i.e. “an act that one does at their own expense that tends to enhance others well-being” [8], trust propensity i.e. “a dispositional willingness to rely on others” [9], and interpersonal trust i.e. “a willingness to accept vulnerability or risk based on expectations regarding another person’s behavior” [10] are fundamental human concepts with implications for personal and societal welfare.

For example, **trust** propensity strongly influences how an individual makes privacy and security decisions, consumes unverified news, and maintains resources in shared online repositories e.g. [11]–[14]. Such scenarios are only likely to grow with the expected growth curves in shared economy, shared augmented reality spaces, and the social internet of things. Hence, understanding and modeling an individual’s trust propensity is an important question for human-centered-computing researchers [15], [16].

Similarly, the humans in such systems may choose to act selfishly or behave **altruistically.** They may choose to provide bandwidth and resources, contribute open-source code, and write Wikipedia entries. In the emerging socio-technical landscape, they may also choose to act differently in the shared economy settings (e.g. Task-rabbit, Uber), and have different preferences for their autonomous cars and bots to behave with others. Further, altruism has been connected with emotional, physical, and financial well-being of individuals as well as communities [8], [17].

Likewise, a person may want to obtain recommendations only from somebody they have high levels of **interpersonal trust**, rent out homes only to somebody they trust, or seek child care services only from somebody they trust. Each of these aspects is already being mediated by mobile phone apps (e.g. Amazon, Airbnb, UrbanSitter) and as the trend is only likely to increase with the emerging internet of things, modeling interpersonal trust is a critical problem for ubiquitous computing research.

Trust, altruism, and interpersonal trust are thus at the core of the design of human-centered computing systems and their importance is only going to increase in the coming years. Hence, understanding them is an important building block in designing well-functioning socio-technical systems and attaining the Internet of People vision which necessitates the creation of a “sociological profile” for mobile phones users that is capable of inferring their behaviors and preferences [18].

Traditional methods for understanding an individual’s behaviors and propensities have been surveys and lab experiments [19]–[21]. Unfortunately, the human-related information taken by observations in restricted, atypical settings involved a limited amount of data that must be contend with various obstacles such as subjective observations, biases, and narrow observation chances while dealing with pressures such as budget, time, and the effort required [22]. Additionally, the reliance on laboratory elicitable features to recognize altruism, trust, and interpersonal trust hinders the progress of the fields of these concepts. Using such labor intensive methods for eliciting the aforementioned concepts essentially prevents scientists from recognizing behavioral features based on mobility or communication traces that range over time and space (e.g. day-night call ratio, diversity of locations visited) to model one’s propensities and behaviors.

This dissertation proposes a new methodology to model altruism, trust, and interpersonal trust based on *phoneotypes* (mobile phone-based observations of a combination of people's traits) [23] that ultimately aim to complement traditional methods like surveys with a ubiquitous data-driven automated method benefitting individuals and communities to take healthier and wiser decisions using their own data.

### 1.2. Related Work

Altruism, trust, and interpersonal trust have been studied across multiple disciplines (e.g., computer science, information science, sociology, psychology, political science, economy) in the past [8], [17], [24]–[29]. In this work, we discuss the related work which is directly connected with the scope of this dissertation i.e. *modeling altruism, trust, and interpersonal trust using phone-based metadata*. Hence, we discuss the related work that clarifies the terminology and suggests different ways to model altruism, trust, and interpersonal trust with a specific focus on their computational modeling. We, also, review some applications and implications of them as well as the recent use of mobile phones to infer different behavioral propensities and traits for individuals.

#### 1.2.1. Trust and Altruism as Concepts

Despite its importance and popularity in various disciplines, a clear scientific definition of trust is not obvious [30]. Not only this, the notions of trust, trust propensity, interpersonal trust, and trustworthiness are often confused [9], [25], [31]. To alleviate such confusion, here we adopt the following definitions for these concepts:

***Trust:*** "the intention to accept vulnerability to a trustee based on positive expectations of his or her actions" [9, p. 909].

***Trust propensity:*** "a dispositional willingness to rely on others" [9, p. 909].

***Interpersonal trust:*** "a willingness to accept vulnerability or risk based on expectations regarding another person's behavior" [10, p. 1]

***Trustworthiness:*** "the willingness of a person B to act favorably towards a person A, when A has placed an implicit or explicit demand or expectation for action on B" [25, p. 65].

While a person's propensity to trust measures their overall willingness to take risks and overall expectations of people to generally behave well, a trustworthy person acts respectfully and with consideration to the needs of other people. Also, interpersonal trust is something specific to a particular relationship between two people. In this dissertation, we focus on trust propensity and interpersonal trust.

Trust is an essential social concept for understanding human behaviors in various fields. The presence of trust preserves many relations and produces much good [25]. For example, trust could allow for the use of low-cost informal agreements rather than expensive complex contracts [26]. In addition, individuals in more trusting communities often feel happier and are more content with life, more involved with their local communities, and have more supportive friends [32]. In computational settings, trust influences purchase patterns in electronic and mobile commerce [33]. Trust is also an important mediator in how individual's deal with security measures, online service agreements, and mobile commerce transactions [11], [34].

Altruism can be defined as "an act that one does at their own expense that tends to enhance others well-being" [8]. To elevate any confusion with a similar concept i.e. "cooperation", altruism and cooperation are two essentially different sociological concepts despite some similarities between them. Although altruism assumes a cost for the benefactor and advantage to the beneficiary, cooperation merely predicates benefit to the beneficiary, the benefactor might also benefit from the transaction [35]. As a consequence, a technical protocol inspired by [36] which considers cooperation would need to keep a ledger of favors given and received between agents (tit for tat behavior), while the one focusing on altruism would simply need to quantify an individual agent's desire to help others.

**1.2.2. Measuring Trust, Altruism, and Interpersonal Trust**

Multiple efforts have attempted to elicit an individual's propensity to trust others and be altruistic [8], [9], [24], [25]. However, previous studies have largely focused on demographic traits (e.g., gender, race) or used lab-based experiments (e.g., Dictator Game, Trust Game) [19], [20]. Using such methods for eliciting trust, interpersonal trust, and altruism often constrains the scope of studies to factors that can be elicited in the lab settings. Thus, there have been very few attempts that have studied the interconnections between long-term, "in the wild", behavioral features based on mobility or communication traces that range over time and space (e.g. day/night call ratios, average travel distance) along with altruism, trust, and interpersonal trust.

**1.2.3. Computational Modeling of Trust, Altruism, and Interpersonal Trust**

Multiple recent efforts have tried to model trust in computational settings. Farrahi & Zia study the *propagation* of trust as a probabilistic stochastic process [37]. Roy et al., propose a pair of complementary measures to determine trust scores of actors in social networks [27] and Zolfaghar & Aghaieb, focus on the evolution of trust in social networks [38]. *However, very little is known about the interconnections between individual trust propensity and phone-based data. Also, there is no previous study that uses deep learning to infer interpersonal trust using ubiquitous data.*

Similarly, modeling altruism has also started receiving some attention in the computational and mobile computing literature. For example, in [39], the authors used attachment transfer theory to understand reciprocal altruism for tourism online shopping using mobile phones. The impact of altruism, topologies, and traffic patterns on mobile social networks have been studied and modeled in [40]. In [41], the authors studied altruism in a delay tolerant network (DTN) based mobile social network application. And finally, in [36] the authors have argued the case for explicitly modeling altruism levels of individuals in peer-to-peer Internet Streaming Broadcast applications.

*However, there are, as yet, no efforts that utilize phone based data to create automated machine-leaning models for individual altruistic propensities.*

### 1.2.4. Altruism, Trust, Interpersonal Trust, and Social Capital

An individual's altruism, trust, and interpersonal trust are often related to their social behavior [42]–[44]. A very important concept in the study of social behavior is that of social capital [45], [46]. In [45], Putnam characterizes social capital as trust, network structures, and norms that promote cooperation among actors within a society for their mutual benefit. He, also, suggests that formal membership, civic participation, social trust, and altruism are indicators of social capital [46]. Such social capital often comes in two variants: bridging and bonding [45]. While bonding social capital is associated with the presence of family and strong personal ties, and provides emotional support, bridging social capital is associated with the presence of acquaintances and weak ties that provide access to newer information and resources. Both of these variants of social capital have been connected with trust and altruism in multiple studies [45], [47]–[51].

Recent Human Computer Interaction (HCI) studies have connected social capital with phone use behavior, thus suggesting that phone use behavior could also be predictive of an individual's trust, altruism, and interpersonal trust [52]–[54]. Trust, altruism, and interpersonal trust have, also, been connected to maintaining inter-personal relationships especially in long distance relationships where face to face interaction is often not possible. Therefore, phone usage patterns could help model individuals' trust, altruism, and interpersonal trust.

### 1.2.5. Using Mobile Phones and Shallow Machine Learning to Understand Humans

Mobile phones (cellphones or smartphones) have become a primary communication device used by billions of people globally. Majority of contemporary mobile phones are equipped with several sensors, and there exists significant literature utilizing mobile phone sensors to automatically infer individual's propensities and personality traits [3], [23], [55], [56]. This dissertation builds upon a recent line of work on *phoneotypic* modeling [23], which defines a *phoneotype* as the "*composite of an individual's traits as observable via a mobile phone*" and argues that a combination of phone-based behavioral features could build a unique signature for an individual which can model facets of the individual's life (e.g. propensity to cooperate).

This was one of the motivations for us to study *phoneotype* associations with altruism, trust, and interpersonal trust. There has been little work on using *phoneotypic,* i.e.

phone-based data to define automated machine-learning approaches for modeling individual altruism and trust propensities and this dissertation seeks to address this gap.

### 1.2.6. Using Mobile Phones and Deep Learning to Understand Humans

There has been a rich array of recent work on modeling human activities using sensors and deep learning [57]–[60]. These efforts range in applications from health to activities of daily living and employ a wide variety of deep learning approaches including Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Autoencoders, Restricted Boltzmann Machines, Restricted Neural Networks (RNN) and Long-Short Term Memory (LSTM). Rather than activity recognition, where the output varies over time, interpersonal relationships are typically modeled over a cumulative time period. This implies that there is only one score to be predicted (and one learning instance) even if the dataset contains one-year worth of human activities. There are no prior works which define deep learning approaches for inferring interpersonal trust and this dissertation tackles this problem.

### 1.2.7. Fairness in Machine Learning Algorithms

We live in an era where various aspects of our lives are determined by computer algorithms. For example, computer algorithms currently help in schools' admission, recruiting, getting loans, and insurances prices [61], [62]. Various recent efforts have shown that computer algorithms might be biased and discriminatory calling for the necessity of creating fairer models [63]. Recently, there are multiple efforts attempting to create fair models across different fields while highlighting the importance of fair models. For example, the authors in [63] assess bias in automated facial analysis algorithms across genders and races, the authors in [64] provide multiple technical solutions to improve fairness in algorithmic decision-making, and the authors in [65] define a mechanism to determine unfairness in classification outcomes for diverse demographics.

Multiple recent efforts in human-centered computing have used mobile and ubiquitous data to infer propensities of individuals (e.g. to cooperate with others) [23], [55]. Often studied under the umbrella of "Reality Mining" which refers to gathering and analyzing sensor-based data related to predicting human behavior and propensities [66]; multiple such efforts have reported high accuracies at the considered prediction tasks.

However, there has been little work at quantifying the "fairness" of such algorithms in terms of how the quality of the predictions varies over different demographic groups (e.g. across gender). For instance, how do the accuracies and the false positive rates vary across genders? This work takes inspiration from the social science grounded approach of "stratified sampling" to define a fairness-aware approach for classifications.

**1.3. Contributions**

In general, this work aims to use mobile phones metadata (Calls, SMS, GPS, Bluetooth) logs*, not contents,* for ethical and privacy considerations, to predict and model human behaviors and propensities. Specifically, this dissertation proposes the following contributions:

(1) To motivate and ground the usage of phone-based features for automatically inferring individual altruism and trust propensities;

(2) To define supervised machine learning models that automatically infer a person's trust propensity;

(3) To define a fair supervised machine learning model that automatically infers a person's trust propensity;

(4) To define supervised and unsupervised machine learning models that automatically infer a person's altruism propensity;

(5) To utilize deep learning to infer interpersonal trust; and

(6) To test, validate, and potentially refine some social science theories related to trust, altruism, and interpersonal trust using 'Big Data' frameworks.

# Chapter 2

# A Fair Approach to Model Trust Propensity Using Supervised Machine Learning

## 2.1. Introduction

Trust is a fundamental human concept that mediates multiple human processes. It facilitates cooperation, supports commerce, and enhances societal well-being [67]. An individual's trust propensity - i.e. "a dispositional willingness to rely on others" - mediates multiple socio-technical systems [9]. Hence, modeling it is very crucial.

Multiple recent efforts have attempted to elicit and model an individual's trust propensity using different methods [9], [25]. Nonetheless, such studies have mostly focused on traits which could be simply observed (e.g., gender, race, age) or elicited in a small period of time in lab settings (e.g., via surveys and game experiments).

Recently, mobile phones along with sensor-based data have been used by multiple researchers to construct rich and individualized models of human behavior in social, spatial, and temporal settings, and link them to individual personality traits and cooperation tendencies [23], [68]–[70].

Given such recent trends and the theoretical literature connecting trust propensity with social capital and social habits such as maintaining interpersonal relationships [42], [43], this work explores the creation of an automated phone-based approach for modeling individual trust propensity. Moreover, this work defines a fairness-aware model using mobile phone metadata to model individual trust propensity trying to avoid bias and discrimination towards gender in the classification process since it has been reported in similar recent works [63] and there has been little work at quantifying the “fairness” of such models.

Such phone-based methods, if successful, could offer low-cost, faster, scalable, and automatic methods for generating insights into trust propensities for millions of users with applications in social computing as well as political systems and sociology. As a consequence, this work investigates the following research questions:

*RQ1: Do long-term phone-use patterns have some associations with an individual’s trust propensity?*

*RQ2: Can a machine learning algorithm be used to automatically infer individual trust propensity based on phone metadata?*

*RQ3: Can a machine learning model be fair in inferring trust propensity?*

In this work, we analyze the data from a ten-week field + lab study to systematically study the interconnections between phone-based behavioral measures (e.g. number of phone calls made) and “ground truth” trust propensity survey scores [21] for 50 individuals. We first discuss a generic prediction approach to infer trust propensity using mobile phone metadata and then focus on a fair prediction approach to enhance fairness in the prediction process.

**2.2. Rutgers Well-being Study (Trust)**

We study the interconnections between trust propensity and phone-based features based on the data gathered as part of Rutgers Well-being Study undertaken at Rutgers. This study was a 10-week field conducted in Spring 2015 and lab study including 59 participants, most of whom were undergraduate students from the aforementioned university.

Initially, all participants were invited to sign consent forms to participate in the study and install an Android app that would record their call, SMS, and GPS logs. Figure 1 shows a screenshot of the app. The app was developed using the "Funf in a box" framework [71] and was released via a URL shared with the study participants.

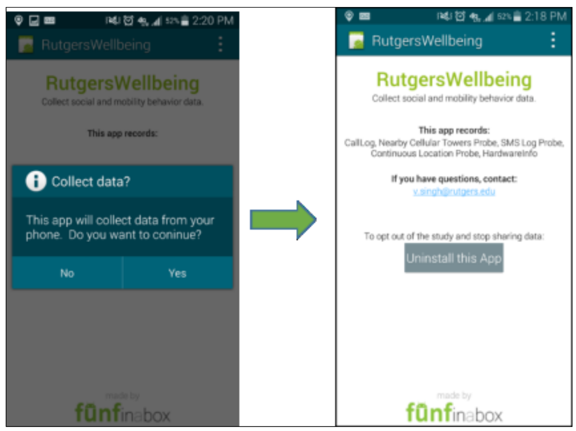


**Figure 1. Screenshot of the Android App.**

The participants were also asked to attend three in-person sessions where they filled out a number of surveys concerning their health, well-being, trust propensity, and some demographics. The order of surveys was randomized for the participants. We use here the trust propensity and demographics surveys for their relevance to this work. There was a compensation of US $20, $30, and $50 respectively for attending the sessions.

Participants' privacy was of utmost priority; hence, anonymized IMEI numbers were used to recognize the participants. All user data were anonymized before analysis. Furthermore, the actual phone numbers or the content of the calls or SMS messages were *not* available to the personnel analyzing and processing the data at any point of time. The permissions required for this study's app (call logs, SMS logs, location logs, and phone identifier information) were intended to be considerably lesser than what is usually required by common apps (e.g. Instagram app on Android). The participation in the study was optional and the participants could withdraw from the study whenever they like. The study was approved by the Institutional Review Board, and all personnel who handled the data in this study were trained and certified in human subject research.

While the study included 59 participants, some of the participants did not complete all the surveys, and some did not enter their unique identifying code consistently across different surveys, resulting in 53 participants. Of these, three participants uploaded location data very rarely (ten or lower instances) - presumably because they turned off location features on their phone - so we removed them from the dataset. This resulted in a dataset involving 50 (32 men, 18 women) participants for whom we have the mobile-based data as well as the scores for the two surveys of interest (more details on surveys presented later). Most participants were in the age group of 18 to 21 years, and the most common education level was “some college”. The median of the participants’ families’ income ranged between US $50,000 to $74,999.

The 50 participants made a total of 25,302 calls with an average of around 506 and a median of 302.5 calls per participant and exchanged 177,263 SMS messages with an average of 3,545 and a median of 2,347 per participant, and visited 14,045 unique locations with an average of about 280 and a median of 295.5 per participant during the period of the study (10 weeks). Table 1 gives a summary of the total, mean, and median for calls, SMS, and GPS locations.

**Table 1. Summary of Calls, SMS, and GPS Location Logs Considered in this Work for Men and Women.**

| Data | Total | | Mean | | Median | |
|---|---|---|---|---|---|---|
| **Calls** | 25,302 | | 506.0 | | 302.5 | |
| **Men/Women** | 17,011 | 8291 | 531.6 | 460.6 | 245 | 326 |
| **SMS** | 177,263 | | 3,545.3 | | 2,347.4 | |
| **Men/Women** | 105,740 | 71,523 | 3304.4 | 3973.5 | 2314.5 | 2518.5 |
| **GPS** | 14,045 | | 280.9 | | 295.5 | |
| **Men/Women** | 8893 | 5152 | 277.9 | 286.2 | 288 | 308 |

### 2.2.1. Trust Propensity Descriptor

The literature discusses several ways of quantifying an individual's trust propensity. For example, games in controlled lab settings such as Trust Game and Dictator Game represent one way of quantifying trust propensities [19], [20]. Surveys that draw individual's behavior in prepared scenarios are another option [21]. Furthermore, a third way is a combination of both: game experiments and lab surveys [26].

In this work, we decided to use a well-known survey "General Trust Scale" to measure trust propensity [21]. The survey has 6 questions whose responses scaled from (5) "Strongly Agree" to (1) "Strongly Disagree" on a five points scale. Some examples of the questions are: "Most people are basically honest" and "Most people are basically good and kind" [21, p. 147]. Besides the prevalent acceptance of the survey (over 2,000 citations as per Google Scholar), we chose this survey as the nature of these questions is not restricted to a specific context and the results could be interpreted in a wide variety of everyday applications. Also, the scale's internal reliability ranges from 0.70 to 0.78 and several studies support its predictive validity [72], [73]. It was developed by selecting items from important trust surveys and has been found to have robust associations with Big Five Personality traits [72], [73].

The scores of the survey are averaged together and normalized as a percentage of the maximum possible score. Thus, the maximum theoretical trust propensity score is 100. In the considered sample, the maximum was found to be 97, the minimum was 40, the mean was 71.5, and the median was 73 as seen in Table 2.

**Table 2. Summary of Trust Propensity Scores for Men and Women.**

| **Minimum** | | **Maximum** | | **Mean** | | **Median** | |
|---|---|---|---|---|---|---|---|
| 40 | | 97 | | 71.5 | | 73 | |
| Men | Women | Men | Women | Men | Women | Men | Women |
| 40 | 43 | 97 | 87 | 72.9 | 69.1 | 73 | 70 |

### 2.2.2. Demographic Descriptors

The participants were surveyed about their demography. Specifically, we obtained the following information: age, gender, marital, level of education (school), and level of family's income.

## 2.3. Mobile Phone Data Features

Trust and socio-mobile behavior have been (indirectly) connected in the past literature in both conceptual and empirical ways. In this work, we consider three major types of socio-mobile features to predict trust propensities.

First, social capital as a concept is connected with both phone use behavior [54] and trust propensities [49], [50]. Hence, we consider a number of phone based features (e.g. number of phone calls, diversity of contacts, and engagement with strong ties) based on the recent literature on using phone meta-data to predict individual social capital or personality traits [23], [54], [56]. In doing so, we do not only consider the frequently used call and SMS metadata, but also consider GPS (location) metadata, which are increasingly being adopted as indicators of physical social activity [74], [75] and also as predictors of an individual's traits and states in their own right [23], [76].

Second, we consider a group of features that have been selected to quantify the trajectories or the mobility behavior of the individuals. These features are related to the concepts of mobility capital (location based analog to social capital) and the notion of a "third place" [77], [78]. (A Third place is a place other than work and home used to build social ties and live a healthy life [78]). Prior research has connected such mobility capital and access to third place with trust [79], [80]. Empirically, these features are based on the recent literature, which has been used to characterize human geo-mobility patterns and study its interconnections with personality and mental health [56], [76].

Third, we consider a set of features that capture the temporal rhythms of human behavior. Conceptually, these features are associated with the notions of circadian rhythms and chronotypes, which have been connected with trust and cooperation in the past literature [81]–[83]. Empirically, these features have been based on recent works that have connected similar features with social capital, cooperation, and well-being [23], [54], [84], [85].

All these features are based on a key working assumption based on Macey and Shneider's model connecting states, traits, and behaviors [86]. Traits are considered to be long-term predispositions, similar to personality attributes. These attributes are often experientially manifested as states, which can be measured indirectly through surveys. States may further manifest through observable and directly measurable behaviors. Hence, here we hypothesize that individual trust propensity traits manifest themselves in the long-term socio-mobile behavior patterns of the users [23], [87]. A summary of the features (N=24) is presented in Table 3.

#### 2.3.1. Social Behavior

##### 2.3.1.1. Social Activity

We quantify the level of social activity as the number of exchanged phone calls, SMS messages, and unique visited locations. A higher count of social activity level suggests an active user and multiple studies have connected individual social activity with social capital and/or trust propensities. High social activity has also been connected with reducing relational uncertainty and as a means of establishing trust in interpersonal relationships [30], [88].

We, also, consider GPS location logs (physical movements) as a proxy of one type of social behavior for it has been used previously to comprehend human social behaviors [23], [75], [76]. The visited locations were updated hourly to balance between getting an idea about the pattern of a user's movement and their phone's battery life. To avoid getting the same amount of locations per participant (24 locations/day), we only count unique locations. The location data were obtained from a mobile phone's GPS as <latitude, longitude> tuple at fourth decimal point resolution, which roughly corresponds to 10m by 10m blocks [23], [89].

$$\text{Social Activity (Call, SMS, GPS)} = \sum \text{Activity}$$

##### 2.4.1.2. Diversity

We are not only considering the total amount of calls, SMS messages, unique locations, but also the diversity (measured as Shannon Entropy) for each one of them, as such a diversity metric has been reported to be associated with multiple personal well-being outcomes and personality traits [69], [90].

$$D_i = -\sum_j p_{ij} \log_b p_{ij}$$

Where $p_{ij}$ is the percentage of social events involving individual 'i' and contact 'j', and 'b' is the total number of such contacts.

**2.4.1.3. Novelty**

The growth of networks plays an important role in social capital [91]. Hence, we, also, consider "new contacts" that are not present in the first four weeks of the data collection period. This feature quantifies how much time users devote to their new contacts as compared to their frequent contacts.

$$\text{Percent New Contacts} = \frac{\sum \text{New Contacts}}{\sum \text{All Contacts}} \text{X } 100$$

**2.4.1.4. Tie Strength**

Previous studies have related strength of ties and trust [92]. Such literature underscores the value of maintaining relationships with both strong and weak ties, and each may yield different types of social capital, and presumably, over periods of time, a propensity to trust others.

Following Williams [93], we connect the concepts of 'bonding' and 'bridging' social capital to those of 'strong' and 'weak' ties as proposed by Granovetter and other researchers [94]–[96]. We conjecture that the relative spread (or concentration) of communication with strong (respectively weak) ties may be a predictor of one's propensity to trust others. It is anticipated that a person would devote at least 33% of their time with their top-third most frequent contacts (proxy for strong ties) [23]. Nonetheless, a high score like 85% may indicate an individual's preference to intentionally engage more with strong ties rather than distributing the communication effort more equally amongst all ties. Hence, we define the following features:

$$\text{Strong Tie Engagement Ratio (STER)} = \frac{\sum \text{communication for highest 1/3 contacts}}{\sum \text{communication}} \text{ X } 100$$

$$\text{Weak Tie Engagement Ratio (WTER)} = \frac{\sum \text{communication for lowest 1/3 contacts}}{\sum \text{communication}} \text{ X } 100$$

### 2.3.2. Spatial Trajectories

Prior research has connected a number of mobility or spatial trajectory related concepts (e.g. mobility capital and access to third place) with trust [79], [80]. Hence, we consider a number of GPS related features to quantify individual behavior.

#### 2.3.2.1. Gyradius

To get a sense about the location distribution of a participant (physical activity), we determine the gyradius (radius of gyration) which is computed as follows. First, we identify the centroid of all the distinct points that a person has visited. Next, we calculate the distance to all points from this center point. The average of such distances traveled is the gyradius [97].

$$\text{Gyradius} = \frac{\sum \text{distance from centroid for each location}}{\text{number of locations visited}}$$

#### 2.4.2.2. Percentage Long-distance Trips

An individual's access to new resources and information is likely to be a function of their access to "far-away" people and places. Hence, we, also, define a feature called Percentage Long-distance trips to quantify the ratio of long distance (above 100 km) trips undertaken by the individual.

$$\text{Percent Longdistance Trips} = \frac{\left|\left|\text{Long Distance Trips}\right|\right|}{\left|\left|\text{All Trips}\right|\right|} \text{X } 100$$

#### 2.4.2.3. Location Loyalty

Location loyalty considers how frequently participants engage with their favorite locations. Past research has connected this loyalty feature with individual well-being [98]. Precisely, we calculate the percentage of time spent in their top three frequented visited locations out of all visited locations.

$$\text{Location Loyalty} = \frac{\sum(\text{time spent in top three locations})}{\sum(\text{time spent in all locations})} \text{X } 100$$

#### 2.4.2.4. Percentage Time Third Place

We, also, introduce here the third place feature which represents the percentage of time spent at the third most visited location by a participant. This is based on the sociological concept of “third place”, proposed by Ray Oldenburg, which states that a person needs a third place – other than work and home (e.g. library, café, worshipping house) – to build social ties and live a healthy life [78]. Past research has connected third places with social capital and trust [80].

$$\text{Percent Time Third Place} = \frac{\sum(\text{time spent in third place})}{\sum(\text{time spent in all places})} \text{X } 100$$

### 2.3.3. Temporal Rhythms

Prior literature has connected circadian cycles, Dark Triad (i.e., narcissism, Machiavellianism, and psychopathy) and trust [81], [82]. The classification of different individual’s chronotype - the tendency for the individual to sleep at a particular time during a day-and/or-night period (24-hour) - has been connected with cheating and Machiavellianism [83].

#### 2.3.3.1. Diurnal Activity Ratio

When we asked some of the participants about their daily activities regarding times when they become productive, and times when they tend to play or sleep (relax), we found that there are two main states: "productive" state from 8 am to 8 pm; "relax" state from 8 pm to 8 am. Hence, to quantify daily patterns of activity and the differences between different phases, we define the following features:

$$\frac{\sum(\text{Call, SMS, Location})\text{when productive}(8\text{am to }8\text{pm})}{\sum(\text{Call, SMS, Location})\text{when relaxed}(8\text{pm to }8\text{am})}$$

#### 2.3.3.2. Weekday/Weekend Activity Ratio

We added another layer of characterization for the abovementioned two states of the daily activity ratio (productive and relaxed) to get more insights out of these circadian rhythms by quantifying the weekdays (Monday to Friday) to weekends (Saturday and Sunday) communication (Call, SMS) ratio.

$$\frac{\sum(\text{Call, SMS})\text{in weekdays}}{\sum(\text{Call, SMS})\text{in weekends}}$$

**Table 3. Summary of *Phoneotypic* (phone-based) Features Defined in this Work.**

| Type | Literature Support | Features |
|---|---|---|
| **Social Behavior Features** | *Conceptual:*<br>Social Capital<br>Putnam [45]; Granovetter [94];<br>Golbeck [88]; Coleman [99];<br>*Empirical:*<br>Eagle et al. [66]; Shmueli et al. [30];<br>Gilbert et al. [95]; deMontjoye et al. [56];<br>Singh & Agarwal [23]; | • Social Activity (Call, SMS, GPS)<br>$\sum \text{Activity}$<br>• Diversity (Call, SMS, GPS)<br>$D_i = -\sum_j pij \log_b p$<br>• Novelty (Call, SMS, GPS)<br>○ Percent New Contacts= $\frac{\sum \text{New Contacts}}{\sum \text{All Contacts}} X\ 100$<br>• Tie Strength (Call, SMS, GPS)<br>○ Strong Tie Engagement Ratio = $\frac{\sum \text{communication for highest 1/3 contacts}}{\sum \text{communication}} X\ 100$<br>○ Weak Tie Engagement Ratio = $\frac{\sum \text{communication for lowest 1/3 contacts}}{\sum \text{communication}} X\ 100$ |

| | | |
|---|---|---|
| **Spatial Trajectory Features** | *Conceptual:*<br>Mobility Capital<br>Golbeck [88]; Coleman [99];<br>Third Place<br>Oldenburg [78];<br>*Empirical:*<br>Pappalardo et al. [100]; Canzian et al. [76]; Singh & Agarwal [23]; Singh et al.[101]; | • Gyradius $= \frac{\sum \text{distance from centroid for each location visited}}{\text{number of locations visited}}$<br>• Percent Longdistance Trips $= \frac{\lvert\lvert\text{Long Distance Trips}\rvert\rvert}{\lvert\lvert\text{All Trips}\rvert\rvert}$ X 100<br>• Location Loyalty $= \frac{\sum \text{(time spent in top three locations)}}{\sum \text{(time spent in all locations)}}$ X 100<br>• Percent Time Third Place $= \frac{\sum\text{(time spent in third place)}}{\sum\text{(time spent in all locations)}}$ X 100 |
| **Temporal Rhythm Features** | *Conceptual:*<br>Circadian Cycles & Chronotypes<br>Jonassona et al. [81]; Lyons & Hughes [82];<br>*Empirical:*<br>Abdullah et al. [85]; Saeb et al. [84]; deMontjoye & Quoidbach [56]; Singh & Ghosh [54]; | • Diurnal Activity Ratio (Call, SMS, GPS)<br>○ DAR $= \frac{\sum \text{Activity when productive}}{\sum \text{Activity when relaxed}}$<br>• Weekday/Weekend Activity Ratio (Call, SMS)<br>○ WWAR $= \frac{\sum \text{(Call,SMS)in weekdays}}{\sum \text{(Call,SMS)in weekends}}$ |

## 2.4. Results

Since multiple applications vary in their requirements of either predicting an exact numeric trust propensity score or working with broader classifications of trust propensity score, we consider both types of applications by undertaking linear regression and classification analyses as follows.

### 2.4.1. Building a Regression Model for Trust Propensity

Here, we first consider predicting trust propensity level as a regression problem; that is, predicting an outcome variable (i.e., trust propensity level) from a set of input predictors (i.e., phone-based features). We use the LASSO (Least Absolute Shrinkage and Selection Operator) regression approach to undertake this [102]. LASSO is a specialized form of regression suitable for scenarios where there are relatively more number of features for a given sample size. It tries to minimize overfitting by penalizing the presence of too many features in the eventual model. It has been applied in similar contexts (in terms of sample size, number of features, and application) in recent human-centered computing research [54], [68]. Similarly, following [54], [68] we evaluate the regression models using the metrics of correlation scores (Cor) (between predicted and actual outcome variables) and the Mean Absolute Error (MAE). While a higher correlation (closer to 1) suggests a higher predictive ability of the considered models, smaller MAE is preferred as it shows that the predictions are closer to the ground truth.

We ran and tested three different regression models: one with the demographic features only, another one with the *phoneotypic* (phone-based) features only, and a third one with a combination of both types of features. The implementation was undertaken using R 3.4.1 [103] and its Lars 1.2 package [104]. To test the statistical significance of these three models, we need an estimate of the (variance) in the effects found. To estimate this, we undertook 100-fold bootstrapping for each LASSO regression model and then undertook unpaired t-tests for the correlation and MAE scores obtained. All comparisons were found to be statistically different at alpha= 0.05 level i.e., Both *>* *Phoneotype* *>* Demography (*>* means statistically significantly higher performance). Table 4 presents the average results for modeling trust propensities using various regression models.

**Table 4. Average Results for Modeling Trust Propensity Using Different Regression Models.**

| Model Type | Cor | SD | MAE | SD |
|---|---|---|---|---|
| **Demography Only** | 0.274 | 0.062 | 9.146 | 0.416 |
| ***Phoneotype* Only** | 0.538 | 0.153 | 7.913 | 1.776 |
| **Both** | 0.544 | 0.153 | 7.711 | 1.541 |

The demography based model obtained on average a correlation of 0.274 (MAE=9.146). The low - but significant - scores for the "demography only" model indicates that the demographic features can explain some (but not a lot) of variance in the trust propensity levels. Phone-based model performed much better with an average correlation score of 0.538 (MAE=7.913).

The combined model using *phoneotype* and demography features performed the best in terms of all metrics and the predicted trust propensity was found to have 0.544 correlation on average with the actual propensity scores (MAE=7.711). This MAE signifies that the predictions are within ±7.711 of the absolute value of the trust propensity scores obtained by the survey (ground truth). Since the trust propensity scores obtained by the survey vary from 40 to 97 as shown in Table 2, ranges of ±7.711 could be considered a reasonable approximation.

Also, we clearly see that the *phoneotype* model and "*phoneotype* + demography" (both) models yield considerably better models than the demography-based model. However, the demographic features were useful in increasing the correlation score for the *phoneotypic* model, thus suggesting that *phoneotypic* features and demographic features are not merely proxies for each other, but rather add newer information when combined.

**2.4.2. Building a Predictive Classification Model for Trust Propensity**

Next, we consider the task of building automated classifiers for trust propensities. In prior research, the same Yamagishi trust scale was used to separate participants into groups of high and low trustors [73]. The survey results predicted behavioral differences between groups of individuals. For instance, groups of high trustors were more likely to cooperate and reciprocate across variations of the prisoner's dilemma and public goods problems [73]. This motivates the analysis in this work on the (phone-based) behavioral differences between high and low trustors and creating computational models for using them in other applications. For instance, an application provider may want to recommend different default privacy settings for individuals with “high” and “low” trust propensity.

Given that there is no universal definition of “high” and “low” trust propensity, we divided the participants into two groups based on the median value (73) for trust propensity survey instrument. The first group (“low” propensity) has 23 participants whose trust score is lower than the median, whereas the second group (“high” propensity) has 27 participants whose trust score is higher than or equal to the median. Similar to the previous analysis, we built three models: one with the demographic features only, another one with the *phoneotypic* features only, and a third one with a combination of both types of features.

We used CfsSubsetEval (Correlation-based Feature Subset Selection) [105] with leave-one-out cross-validation in Weka 3.8.1 [106], [107] which ranks the best subset of the 24 features described previously by determining the predictive capability of each feature in company with the degree of redundancy between them. The best subsets of features are correlated with the target variable and have low intercorrelation [105]. We found that the best subsets of features in most of the folds are the ones shown in Table 5.

**Table 5. Selected Features for Different Prediction Models.**

| | |
|---|---|
| **Demography Only** | Age, School (education level) |
| ***Phoneotype* Only** | SMS Entropy, Weekday Weekend Call Ratio, Percent Time Third place, Percent Long Distance Trips |
| **Both** | SMS Entropy, Weekday Weekend Call Ratio, Percent Time Third Place, Percent Long Distance Trips, Age |

To define and test a machine learning based classifier whose *phoneotypic* features can statistically significantly improve the ability of predicting trust propensity when compared to the demographic features, we took 10-fold cross-validation and repeated it 10 times to get 100 different values for CA, AUC, and F1 and build the predictive models. AUC stands for (Area Under the Receiver Operating Characteristic Curve), CA means classification accuracy and F1 score represents the harmonic mean between precision and recall [108], [109]. The aforementioned features were used to test out three well-known machine learning algorithms for classification. Specifically, we used Adaptive Boosting (AdaBoost), Random Forest, and KStar. We also used a Zero-R model which simply classifies all the instances into the majority class, as a baseline to help interpreting the performance of the considered models. Statistical comparison was undertaken using unpaired t-tests (at alpha= 0.05 level) suggesting that for AUC, CA, and F1: *Phoneotype* *>* Demography, Both *>* Demography, Both (not significantly different from) *Phoneotype*; (*>* means statistically significantly higher performance). All three models above were significantly better than Zero-R.

**Table 6. Average Results of Predicting Trust Propensity Using Different Classification Methods.**

| Method | Demography Only | | | *Phoneotype* Only | | | Both | | |
|---|---|---|---|---|---|---|---|---|---|
| | *AUC* | *CA* | *F1* | *AUC* | *CA* | *F1* | *AUC* | *CA* | *F1* |
| **AdaBoost** | **0.68** | **0.69** | **0.66** | **0.81** | 0.76 | **0.75** | **0.83** | **0.79** | **0.78** |
| **Random Forest** | 0.58 | 0.62 | 0.60 | 0.79 | **0.77** | **0.75** | 0.82 | 0.78 | 0.77 |
| **K Star** | 0.63 | 0.62 | 0.60 | 0.73 | 0.64 | 0.62 | 0.80 | 0.73 | 0.72 |
| **Zero-R** | 0.50 | 0.54 | 0.38 | 0.50 | 0.54 | 0.38 | 0.50 | 0.54 | 0.38 |

Table 6 shows that the demography-based model returned the best CA of 69%, AUC of 0.68, and F1 score of 0.66. The *phoneotype*-based model yielded a better classification performance and the best CA was 77%, AUC was 0.81, and F1 was 0.75. While the demographic features contained some predictive power, we observe that *phoneotypic* models considerably outperform demographic models.

It is also clear that the *phoneotypic* model outperformed the Zero-R model. The *phoneotypic* model performed 62% better than the Zero-R model in terms of AUC, 42.6% better in terms of CA, and 97.4% better in terms of F1.

We also considered the cases where the demographic data may be available to the phone app. In such a case, the combined model (demography + *phoneotype* data) yielded an even higher performance with a CA of 79%, AUC of 0.83, and F1 of 0.78.

Hence, we note that a phone-features based model beats baseline majority classification and also goes beyond static demographic descriptors (e.g. age, gender, education) for predicting trust propensities. This underscores the potential for using phone-based (*phoneotypic*) features to build automatic classifiers for individual trust propensities. One way to interpret these results is that having mobile sensing data for 10 weeks may allow for the creation of a detailed model for personal behavior based on the aforementioned idea of phone behavior being akin to a vast psychological questionnaire, being constantly filled out [7].

#### 2.4.3. FAIRSTART: A Fair Approach to Model Trust Propensity

To fairly predict trust propensity, we have identified the following characteristics for a fair method to use. First, it should not use demographic attributes to make predictions. Next, it should be amenable to small datasets. Hence, we choose to undertake validation based on leave-one-out cross-validation, which tries to balance the learning opportunities with the testing rigor. Lastly, we use balanced stratified sampling to make sure we get a random but equal number of instances of each considered demographic group (e.g. men, women) to maintain fairness of representation at the input level [110]. Also, we use multiple decision trees that will form a random forest as classifiers for random forests are good at reducing bias [111].

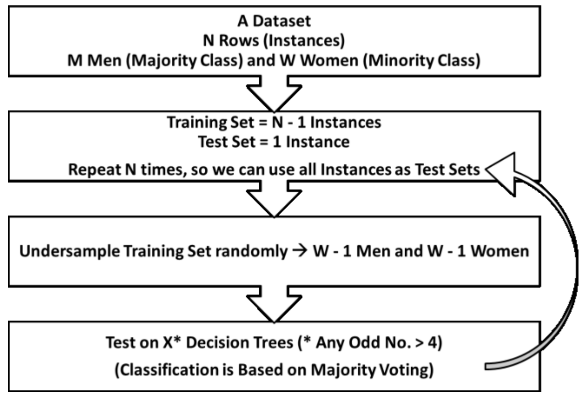


**Figure 2. Flowchart Explaining Our Method (FAIRSTART).**

FAIRSTART (**Fair**ness-aware **str**a**t**ified **r**andom fores**t**) is designed as follows: given a dataset with N instances (rows), M of which belongs to the majority class of the sensitive attribute (men here), and W of which belongs to the minority class of the sensitive attribute (women here), we split the dataset into (N - 1) instances for training and we leave 1 instance for testing (the ultimate goal is to do a leave-one-out cross-validation). Then, we undersample the training set randomly to have equal number (W - 1) for the two classes. Next, we pass the set to X* decision trees for classification (* any odd number of decision tress greater than 4 to facilitate the majority voting process); the X decision trees collectively form a random forest whose final classification result is based on majority voting. We repeat this process N times to make sure that we use all N instances in the test sets. Figure 2 illustrates the proposed method.

In this section, we concentrate on fair classification (prediction) of trust propensity using mobile phone metadata related to gender discrimination whether or not men and women are equally and justly treated. To quantify fairness, we use two well-accepted fairness definitions commonly used in a binary classification setting similar to the dataset here: *equal opportunity* and *demographic parity* [112], [113]. *Demographic parity* requires that a binary decision (e.g. trust propensity class) must be independent of the protected (sensitive) attribute (e.g. gender). In other words, if $\hat{Y}$ is a binary decision $\hat{Y} \in \{0,1\}$, and a binary protected attribute $A \in \{0,1\}$, we want to satisfy the following condition:

$$\Pr\{\hat{Y} = 1 \mid A = 0\} = \Pr\{\hat{Y} = 1 \mid A = 1\}.$$

In a binary decision $\hat{Y} \in \{0,1\}$, *Equal opportunity* prefers one outcome $\hat{Y} = 1$ (e.g. high trust propensity) and mandates non-discrimination merely to it. In other words, we want to satisfy the following condition:

$$\Pr\{\hat{Y} = 1 \mid A = 0, Y = 1\} = \Pr\{\hat{Y} = 1 \mid A = 1, Y = 1\}.$$

*Demographic parity* requires the selection of at least 80% of any gender of the rate for the gender with the highest rate, while *equal opportunity* helps in fixing two well-known flaws with *demographic parity* [112], [113].

We used Scikit-learn [114] to build the models based on the proposed method. We built two models: a baseline model and a FAIRSTART model. The FAIRSTART model is based on our proposed method as explained earlier (with 11 decision trees). The baseline model is based on 11 standard Scikit-learn decision trees without any sampling and with leave-one-out cross-validation. We repeated the experiment 10 times to obtain stable averages for results.

Table 7 shows that the baseline model's accuracy is 66.80% on average. Further, there is a significant difference in the accuracy levels for men (62.50%) and women (74.44%).

Hence, there is indeed a noticeable difference in the performance of the classifier for different demographic groups.

**Table 7. Average Results of Predicting Trust Propensity Using FAIRSTART Model.**

| Model | Total Accuracy | Accuracy-Men | Accuracy-Women | Equal Opportunity | Parity Difference |
|---|---|---|---|---|---|
| **Baseline** | 66.80 | 62.50 | 74.44 | 0.26 | 0.07 |
| **FAIRSTART** | 75.60 | 72.81 | 80.56 | 0.18 | 0.04 |

The proposed FAIRSTART model's accuracy is 75.60% on average that is 13.17% better in relative terms than the baseline model while being fairer by better satisfying both fairness definitions (less is better). Statistical comparison was undertaken using two-tailed unpaired t-tests (at alpha= 0.05 level) suggesting that for Accuracy, Accuracy-Men, Accuracy-Women, and Equal Opportunity, the FAIRSTART model was statistically significantly better than the baseline model. At the same time, the change in the Parity Difference was not statistically significant (it is partially significant in a one-tailed unpaired t-test (at alpha= 0.10 level)). Nevertheless, it still followed a trend of improvement using FAIRSTART model compared to the baseline model.

### 2.4.4. Behavioral Features Associated with Trust Propensity

Besides creating automated methods for identifying an individual's trust propensity levels, one of the goals of this work is to understand the socio-mobile behavior of individuals with different propensities to trust. Thus, we undertook a *post-hoc* Pearson's correlation analysis between trust propensity scores and the *phoneotypic* features. In the interest of space, we only report the correlations that were found to be (at least marginally i.e., $p<0.10$) significant in Table 8.

**Table 8. Correlation between Phone-based Features and Trust Propensity. (** 0.01, * 0.05, ° 0.10).**

| Feature | Pearson's Correlation | p-value |
|---|---|---|
| **Social Activity (Call)** | +0.237° | 0.097 |
| **Strong Tie Engagement Ratio (Call)** | +0.249° | 0.081 |
| **Weekday Weekend Ratio (Call)** | +0.371** | 0.008 |
| **Gyradius** | -0.271° | 0.057 |
| **Percent Time Third Place** | -0.252° | 0.078 |

We note that people who have high trust propensity tend to be more socially active, yet tend to limit or concentrate their social activities both spatially and temporally. For instance, individuals with higher trust propensity tend to call more often ($r$= +0.237). This can be understood as trust propensity being associated with healthy social relationships and higher call activity captures such behavior [26], [32].

Next, we notice that individuals with higher trust propensity tend to have higher preference for *concentrating* social activities in multiple ways. First, they show a marked preference for engaging in phone calls with their "strong ties" as opposed to spending it equitably with all contacts ($r$= +0.249). The notion of concentrating social activities continues temporally and we notice that the individuals with higher trust propensity tend to concentrate their calling more over the weekdays as opposed to spreading it evenly across all days of the week ($r$= +0.371).

This aspect of concentrating activities becomes even more prominent when we consider their spatial trajectories. Individuals with higher trust propensity tend to have a smaller gyradius ($r$= -0.271) and spend less time at even their third-favorite place ($r$= -0.252), presumably preferring to spend time at their top two favorite locations. One way to interpret these results is that those who travel further and frequently tend to have limited chances to build strong ties and the lack of strong ties has been associated with a lower trust propensity in the past [115].

**2.5. Discussion**

*The first research question (RQ1) for this work was: Do long-term phone-use patterns have some associations with an individual's trust propensity?*

The Pearson's correlation analysis in the preceding section indicates that multiple phone-based features are correlated with an individual's trust propensity. We notice that the individual effect sizes are small and the p-values for multiple of the associations are considered marginally significant. We acknowledge this as a limitation of the sample size (50 participants), but our confidence is increased by considering that many of the same features show up to be prominent in the features selected by LASSO regression and those selected by the classification algorithms.

Hence, while further testing on individual features is needed as part of the future work, the exploratory work here suggests multiple associations between trust propensity and phone-based social behavior.

The individual associations found can also be connected with the literature connecting trust and social relationships. First, the findings suggest that trust propensity builds more on "strong ties" rather than "weak ties" [94]. While, higher social activity was positively associated with trust propensity, it was also found to grow in concentrated (social, spatial, temporal) accumulation of such connections. Presumably, repeated social interactions with familiar faces and places, i.e. "bonding" social capital is conducive for developing trust propensities. Conversely, it is possible that those with higher trust propensities tend to build and focus on a small number of relationships.

According to the social identity and self-categorization theories, group-based stereotypes or in-group favoring behaviors might explain how an individual trusts strangers [116]. While individuals normally have good expectations on strangers (out-group members), they anticipate a better treatment when it comes to in-group members (in-group favoritism) which eventually transforms into a greater trust propensity to an in-group, not an out-group member [116]–[118]. Constant interactions with such in-group members may result in a longer-term internalization of this trust propensity. All of these aspects are associated with the positive association observed between concentrating social activities - socially and geographically - and a higher trust propensity.

The relational uncertainty theory (RUT), which studies the degree of confidence people have in their perceptions of involvement within interpersonal relationships [119] gives yet another perspective to understand the results. It suggests that trust in long distance relationships is negatively associated with relational uncertainty and reducing uncertainty via constant communication (Social Activity (Call), Strong Tie Engagement Ratio) might be positively associated with trust building. While RUT has mostly been studied in terms of face to face interactions in the past, the current results suggest that similar relationships might hold over phone interactions too.

*The second research question (RQ2) for this work was: Can a machine learning algorithm be used to automatically infer individual trust propensity based on phone metadata?*

The three types of analysis adopted in this work (regression analysis, correlation analysis, and the classification models) suggest that machine learning and in general analytics approaches can indeed be used to infer individual trust propensity based on phone metadata to a large extent. The regression analysis can estimate the individual trust propensities with high correlation (0.544) and within a margin of ±7.711 over a range of 40 to 97. Complementing phone features with demographic data, where available, could yield even better performance. For instance, the classification analysis yielded up to 79% accuracy (AUC=0.83; F1=0.78) based on such models.

Given the modest sample size, we concentrate on finding general patterns and trends over the three analysis techniques here. We can see a consistency in the results across the three analysis methods suggesting that socio-mobile signals as observed via a phone (*phoneotype*) could indeed be used to infer trust propensity of an individual to a reasonable extent.

*The third research question (RQ3) was: Can a machine learning model be fair in inferring trust propensity?*

The proposed FAIRSTART model's accuracy was 75.60% on average i.e. 13.17% better in relative terms than the baseline model whose accuracy was 66.80% while being fairer by better satisfying both fairness definitions: Equal Opportunity and Parity Difference. Further, there was a significant difference in the accuracy levels for men (62.50%) and women (74.44%). Therefore, there is indeed a noticeable difference in the performance of the classifier for different demographic groups.

One unexpected observation in the results was the higher accuracy for women (the minority class for the sensitive attribute) despite there being more learning instances for men. This finding points to the difference between less raw data and fewer learning instances, which would be important for fairness aware app designers working in similar areas. One of the important discussion points in fairness aware machine learning is the idea that the prediction quality is often worse for the demographic minorities because the datasets contain lesser instances of people with those attributes. Hence, techniques like data augmentation - creating artificial samples for the minority class - have been proposed to counter this issue [120]. However, such an approach confounds fewer instances with lesser information about that group of individuals.

In reality mining scenarios, since many raw data points are processed into one attribute for a person, sometimes the individuals in the minority demographic class may actually have better quality of data. For example, in the considered dataset here, women form the demographic minority (18 out of 50 people), nevertheless, women record higher number of phone calls, SMS messages, and unique GPS locations compared to men as presented in Table 1. Hence, it would not be surprising to find that Accuracy-Women to be more accurate than Accuracy-Men as presented in Table 7.

This difference between demographic minority class and lesser learning data condition could shape interesting methodological and ideological discussion in the human-centric computing. In particular, we would expect the notions of minority class to expand to include those with lesser supporting raw data or even the social characteristics of the individuals like less social activity, or introversion, or lesser familiarity with technology.

### 2.5.1. Privacy of User Data and Ethical Considerations

All data used in this study were hashed and anonymized as discussed in the study design. The permissions needed for the study (call logs, SMS logs, location logs, and phone identifier information) were designed to be significantly lesser than those typically adopted by popular apps. Lastly, the participation in the study was on a voluntary basis, and the participants could drop out at any time.

We also note the ethical concerns surrounding assigning an individual a score based on their propensity to trust. While such scores could be used by an individual to receive recommendations for privacy, social networking, and mobile commerce applications, they could also be used by commercial and other organizations to infer individual trust propensities. Similar concerns have been raised about the traditional paper survey based methods administered by any organization, and also newer automated techniques that use social media and phone data to assign health, well-being, or similar “suitability” scores to individuals [121]. Instead of shunning away from reporting such results, or shrouding such research in secrecy, we adopt the approach of raising awareness about these new possibilities and informing the policy debate surrounding them.

### 2.5.2. Limitations

This study has some limitations. First, we acknowledge that the analysis in this work focused only on correlations and it does not imply causation. Next, the homogeneity of the sample (most of the participants were undergraduate students from the same university) stops us from generalizing the findings to larger populations, yet the homogeneity permits isolating socio-mobile behavior as a predictor.

From a methodological perspective, we note the multiple comparisons undertaken in the correlation analysis. While such multiple comparisons are often “corrected” using Bonferroni or Bonferroni-Holm correction to maintain the confidence in the associations found, we do not do so in this work because the analysis undertaken here is *posthoc* and intended to help interpret the observed prediction results rather than being prescriptive in its own right. Similarly, we acknowledge the issues associated with the use of a relatively large number (24) of possibly collinear features in regression given the modest sample size (50). While this makes the interpretation of individual feature coefficients difficult, the model’s average correlation scores of 0.538 for *phoneotype* (respectively 0.544 for *phoneotype* + demography) remain interpretable, especially given the use of LASSO regression, which is purposely designed to handle such scenarios [102].

While we consider the results in this work to be exploratory, the results from the regression analysis, correlation analysis, and the classification models point to a common theme that there are indeed interconnections between phone-based behavioral features and individual trust propensities. These results motivate further work in this direction to expand the understanding of the associations between socio-mobile behavioral data and trust propensity.

### 2.5.3. Implications

With further validation, this line of research could have multiple implications for individuals as well as the society.

We suggest the use of such methodologies to be based on opt-in. The participants who opt-in to such automated trust propensity scoring apps could get better customized recommendations for privacy, security, social networking, news, and mobile commerce applications. For instance, in [122], the authors found that the propensity to trust is an antecedent of the attitudes of mobile users toward in-app advertisements. Similarly, understanding trust propensity is likely to be the most relevant trust antecedent in contexts involving unfamiliar actors [9]. This is important to understand societal changes as well as emerging socio-technical contexts like the sharing economy [123]. Generally, the suggested phone-based method here could open ways to better model human beings based on ubiquitous sensing.

At a societal level, such applications could alleviate the need to run costly annual surveys to access the *trust-based* “state of the nation” as proposed in [26]. Instead, automated methods could be used to create a real-time nation-wide trust propensity census and make it part of the public policy and decision making process. Further, an ability to study the phenomenon of trust propensity and its “in the wild” dynamics at scale could substantially advance the literature in multiple fields (e.g. economics, psychology, management) that study trust and trust propensity. For instance, this approach could help the researchers in many fields to ask research questions that were simply not feasible in lab-based settings (e.g. contagion in trust propensities across networks of millions of individuals).

Mobile apps today mediate multiple human functions ranging from mental health prediction to access to better jobs, friends, and information. For instance, ACLU has recently sued Facebook for showing ads for technical jobs to younger men than other demographics [124]. We can expect similar issues to become even more important in areas like trust, which mediate processes like renting of houses, cars, and services. Hence, approaches for making them fairer are important for the human-centered computing research. This work identifies some of the characteristics which are important for a family of applications in human-centered reality mining and defines a new social-science literature grounded approach for making the algorithms fairer. This work hence argues a case for expanding the discussion on “value sensitive design” [125] to include the design of machine learning algorithms in human-centered machine learning.

# Chapter 3

# Modeling Altruism Propensity Using Supervised and Unsupervised Machine Learning

## 3.1. Introduction

In this work, *we define altruism as any act (or behavior) one does at their own expense that tends to enhance others well-being* [8]. There have been multiple efforts aimed at trying to elicit an individual's propensity to behave altruistically with others [8], [24]. However, previous works have generally focused on traits that could be simply observed (e.g., gender, ethnicity, age) or elicited in a small period in lab settings (e.g., via surveys and games).

Lately, mobile phones along with sensor-based data have been used by scientists to construct rich and individualized models of human behavior in social, spatial, and temporal settings and link them to depression, happiness evaluations, and school GPA (Grade Point Average) [30], [68], [126]. This progress motivates utilizing phone-based models for predicting altruism too. Such a phone-based method, if successful at predicting altruism propensity, may offer a low-cost, faster, scalable, and automatic process for making insights into altruism levels of billions of users in the big data era.

Hence, in this work, we systematically identify the associations between phone-based behavioral indicators and altruism and quantify the predictive power of such personal big data in inferring a person's propensity to behave altruistically.

The main contributions of this work are three-fold:

(1) To motivate and ground the usage of phone-based features for inferring altruism propensity;

(2) To identify the associations between long-term “in the wild” socio-mobile behavior and altruism propensity; and

(3) To define a machine learning model that automatically infers a person’s propensity to be altruistic.

In this chapter conversely to the previous one, we use “k-Means++” [127] (an unsupervised machine learning algorithm) to group the participants into naturally occurring clusters/categories based on their altruism score before evaluating phone data-based models to infer the right altruism category for individuals instead of a split around the median as such a median splitting method has limitations in terms of its ability to capture the underlying dynamics of the data due to the arbitrary split point selection. In other words, median split is variable-oriented, not people-oriented [128]–[130].

**3.2. Rutgers Well-being Study (Altruism)**

We study the interconnections between altruism and phone-based features on the data collected as part of the Rutgers Well-being Study. This study was a ten-week field and lab study conducted in Spring 2015 including 55 participants, most of whom were undergraduate students from Rutgers, The State University of New Jersey.

Initially, all participants were invited to sign a consent agreement to participate in the study and install an Android mobile app. The mobile app could record their call, SMS, and GPS logs, not content. The participants were requested to be present in-person for three sessions where they filled out a number of surveys concerning their health, well-being, altruism, and some demographics. There was a compensation of $20, $30, and

$50 for attending the sessions. An approach of increasing the compensation for each session was adopted in an effort to reduce the dropout rate over the ten-week period of the study. We use here the altruism and demographics surveys for their relevance to this work.

Participants' privacy was an utmost priority; hence, anonymized IMEI numbers were used to recognize the participants. Also, the dataset was hashed before analysis. The participation in the study was voluntary and the participants could withdraw from the study at any time. All staff who handled the data in this study were trained and certified in human subject research.

The participants' ages varied from 18 to 21 years. Of these, 35 were men and 20 were women. Most of the participants were single and the median of their families' income ranges between US $50,000 to $74,999. Altruism propensities (tendencies) were quantified using a survey (details follow), whereas the phone-based features have been attained from an app installed in their Android mobile phones. The app was developed using the "Funf in a box" framework [71]. We decided not to upload this app to Google Play Store to make sure that no one beside the participants has an access to the app or its data. The goal of this work is to test the feasibility of eventually replacing such in-lab surveys with automated phone-based methods.

The 55 participants made a total of 28,132 calls with an average of about 511 and a median of 312 calls per participant and exchanged 187,720 SMS messages with an average of 3,413 approximately and a median of 2,423 per participant, and visited 14,905 unique locations with an average of 271 and a median of 284 per participant during the period of the study (10 weeks). Table 9 gives a summary of the total, mean, and median for calls, SMS, and locations.

**Table 9. Summary of Calls, SMS, and Locations in this Work.**

| Feature | Total | Mean | Median |
|---|---|---|---|
| **Calls** | 28,132 | 511 | 312 |
| **SMS** | 187,720 | 3,413 | 2423 |
| **Unique Locations** | 14,905 | 271 | 284 |

### 3.2.1. Altruism Descriptor

In this work, we decided to use a well-known survey to measure altruism: "The Self-report Altruism Scale" (SRA) by Rushton et al. [131]. The survey has 20 questions whose responses scaled from (5) "very often" to (1) "never" on a five points scale. Examples of the questions are: *"I have given money to a charity*" and "*I have pointed out a clerk's error (in a bank, at the supermarket) in undercharging me for an item*" [131]. Besides the widespread adoption of the survey (over 800 citations as per Google Scholar), we chose this survey as the nature of these questions is not restricted to a specific context and the results could be interpreted in a wide variety of everyday applications. Also, SRA has an adequate validity correlations with related measures and a high reliability of $\alpha = 0.80$ [132].

Since the survey has 20 questions worth 5 points each, the maximum theoretical altruism score is 100. In the considered sample, the maximum is found to be 95, the minimum is 31, the median is 50, and the mean is 54.07 as shown in Table 10.

**Table 10. Summary of Altruism Propensity Scores.**

| Minimum | Maximum | Median | Mean |
|---|---|---|---|
| 31 | 95 | 50 | 54.07 |

### 3.2.2. Demographic Descriptors

The participants were surveyed about their demography. We collected the following information: age, gender, marital status, race, level of education (school), and level of family's income.

## 3.3. Mobile Phone Data Features

To come up with a good representation of an individual's social-mobile behavior, we surveyed the related literature which focuses on connecting phone behavior with individual behaviors and social outcomes (e.g., [30], [36], [39], [55]). For example, social capital as a concept is connected with both phone use behavior [54] and altruism [51]. Social capital often comes in two variants: bridging and bonding [45]. Hence, we link the concepts of weak and strong ties to bridging and bonding social capital to predict one's propensity to altruism [23], [93], [94]. We use call and SMS logs to represent the features that carry "social traits" concepts for mobility and altruism and their interconnections [133]. We, also, consider location logs (physical movements) as a proxy of social behavior for it has been used previously to comprehend human social behaviors [75] and human social and geo-spatial behavior are inherently connected with each other [74].

Based on the Call, SMS, location data collected from the app, we define the following set of features:

### 3.3.1. Level of Social Activity

Level of Social Activity represents the activity of a user as obtained through counting exchanged phone calls, messages, and unique visited locations. A higher count of social activity level suggests an active user. The visited locations were updated hourly to balance between getting an idea about the pattern of a user's movement and their phones' battery life. To avoid getting the same amount of locations per participant (24 locations/day), we focus on unique locations. The location data were gained from a mobile phone's GPS as <latitude, longitude> tuple at fourth decimal point resolution, which roughly corresponds to 10m by 10m blocks [23], [89]. We are not only considering the total amount of calls, but also the total durations of such calls because they are related to social activity. We assume that a person who makes or receives (I/O) numerous long calls may have more social life and this may be associated with being more altruistic [8]. Thus, we consider the following features:

$$\text{Social Activity (Call, SMS, Location)} = \sum \text{Activity}$$

$$\text{Total Call Duration} = \sum \text{Time Spent on I/O calls}$$

**3.3.2. Diversity (Calls, SMS, Location)**

We are not merely considering quantifying calls, SMS messages, unique locations, but also the diversity (measured as Shannon Entropy) for each one of them, as such a diversity metric has been reported to be associated with various personal well-being outcomes and personality traits [69], [90].

$$\text{Diversity (Call, SMS, Location):}$$

$$\mathrm{D_i} = -\sum\nolimits_j \mathrm{p_{ij}} \log_b \mathrm{p_{ij}}$$

Where $p_{ij}$ is the percentage of social events involving individual 'i' and contact 'j', and 'b' is the total number of such contacts.

### 3.3.3. Novelty (Call, SMS, Location)

The growth of networks plays an important role in social capital [91]. Hence, we, also, consider "new contacts" that are not present in the first four weeks of the data collection period. This feature quantifies how much time users devote to their new contacts as compared to their frequent contacts.

$$\text{Novelty (Call, SMS, Location):}$$

$$\text{Percent New Contacts} = \frac{\sum \text{New Contacts}}{\sum \text{All Contacts}} \text{ X } 100$$

### 3.3.4. Reciprocity (Call, SMS)

Besides the frequency of communication, the ease with which communication is conducted is also an important property of an individual's social behavior. We anticipate approachability of individuals to be associated with their civic participation and social capital levels [54]. Such social capital levels have been associated with altruism [51]. Hence, we compute the ratio of incoming to outgoing calls and SMS text messages and also the percentage of missed calls as follows.

$$\text{In Out Ratio (Call, SMS):}$$

$$\text{IOR} = \frac{\text{Incoming communication count}}{\text{Outgoing communication count}}$$

$$\text{Missed Call Percentage} = \frac{\sum \text{ missed calls}}{\sum \text{calls}} \text{ X } 100$$

### 3.3.5. Strong and Weak Ties Engagement Ratio (Call, SMS, Location)

It is anticipated that a person would devote at least 33% of their time with their top third contacts. Nevertheless, a higher score like 80% may indicate an individual's preference to pointedly engage more with strong ties rather than spreading the communication effort more equally among all ties. We were inspired by prior studies linking strength of ties and altruism [134], [92], and conjecture that the relative spread (or concentration) of communication with such strong ties may be a predictor of one's propensity to be altruistic.

$$\text{STR}= \frac{\sum \text{communication for highest} \left(\frac{1}{3}\right) \text{contacts}}{\sum \text{communication}} \text{ X } 100$$

$$\text{WTR}= \frac{\sum \text{communication for lowest} \left(\frac{1}{3}\right) \text{contacts}}{\sum \text{communication}} \text{ X } 100$$

**3.3.6. Temporal Rhythms (Call, SMS, Location)**

Prior literature has connected animal rhythms and circadian cycles and altruism [81]. The characterization of different individual's chronotype - the tendency for the individual to sleep at a particular time during a 24-hour period - colloquially "morningness" or "eveningness" has been connected with cheating and Machiavellianism [83]. When we asked some of the participants (mainly students) about their daily activities, times when they become productive, and times when they tend to play or sleep (relax), we found that there are two main states: "productive" state from 8 am to 8 pm; "relax" state from 8 pm to 8 am.

$$\frac{\sum(\text{Call, SMS, Location})\text{when productive}(\text{8am to 8pm})}{\sum(\text{Call, SMS, Location})\text{when relaxed}(\text{8pm to 8am})}$$

We added another layer of characterization for the abovementioned two states of the daily activity ratio (productive and relaxed) to get more insights out of these circadian rhythms by quantifying the weekdays (Monday to Friday) to weekends (Saturday and Sunday) communication (Call, SMS) ratio.

$$\frac{\sum(\text{Call, SMS})\text{in weekdays}}{\sum(\text{Call, SMS})\text{in weekends}}$$

Table 11 summarizes all the (N=24) *phoneotypic* i.e. phone-based behavioral features in this work.

**Table 11. Summary of *Phoneotypic* Features in this Work.**

| Feature | Definition |
|---|---|
| **Level of Social Activity** | Social Activity (Call, SMS, Location)<br>$\sum$ Activity<br>Total Call Duration $= \sum$ Time Spent on I/O calls |
| **Diversity** | Diversity (Call, SMS, Location):<br>$D_i = -\sum_j p_{ij} \log_b p_{ij}$ |
| **Novelty** | Novelty (Call, SMS, Location):<br>Percent New Contacts= $\frac{\sum \text{New Contacts}}{\sum \text{All Contacts}}$ X 100 |
| **Reciprocity** | $\text{IOR} = \frac{\text{Incoming communication count}}{\text{Outgoing communication count}}$<br>Missed Call Percentage= $\frac{\sum \text{ missed calls}}{\sum \text{calls}}$ X 100 |
| **Strong and Weak Ties Engagement Ratio** | STR= $\frac{\sum\text{communication for highest}\left(\frac{1}{3}\right)\text{contacts}}{\sum \text{communication}}$ X 100<br>WTR= $\frac{\sum\text{communication for lowest}\left(\frac{1}{3}\right)\text{ contacts}}{\sum \text{ communication}}$ X 100 |
| **Temporal Rhythms** | $\frac{\sum(\text{Call, SMS, Location})\text{when productive}(8\text{am to }8\text{pm})}{\sum(\text{Call, SMS, Location})\text{when relaxed}(8\text{pm to }8\text{am})}$<br>$\frac{\sum(\text{Call, SMS})\text{in weekdays}}{\sum(\text{Call, SMS})\text{in weekends}}$ |

## 3.4. Results

Multiple applications vary in their requirements of either estimating an exact numeric for altruism score (e.g. for studying altruism levels in social science studies) or working with broader classifications of altruism score (e.g. for suggesting different default preferences for bandwidth sharing). Hence, we consider both types of applications by undertaking linear regression and classification analyses as follows.

### 3.4.1. Building a Regression Model for Altruism

We first consider predicting altruism level as a regression problem; that is, predicting an outcome variable (i.e., altruism level) from a set of input predictors (i.e., phone-based features). We use the LASSO (Least Absolute Shrinkage and Selection Operator) regression approach to undertake this process [102]. It has been applied in similar contexts (in terms of sample size, number of features, and application) in recent human-centered/ubiquitous computing research [68], [54]. Similarly, following [68], [54], we assess the regression models using the metrics of correlation scores (between predicted and actual outcome variables), the Root Mean Square Error (RMSE), and the Mean Absolute Error (MAE). While a higher correlation (closer to 1) suggests a higher predictive ability of the considered models, smaller RMSE and MAE are preferred as they show that the predictions are closer to the ground truth altruism survey.

We ran and tested three different regression models: one with the demographic features only, another one with the *phoneotypic* (phone-based) features only, and a third one with a combination of both types of features. All demography features were found to be significant (Age, Gender, School, Race, and Income) in the demography only model except “Marital”. All *phoneotypic* features were found to be significant (except Weak Ties (Location)) in the *phoneotype* only model. Finally, all demography and *phoneotype* features were significant (except In Out ratio (Call), Weak Ties (Location), and Race) in the combined model. The implementation was undertaken using R 3.4.1 [103] and its Lars 1.2 package [104]. Table 12 presents the results for the evaluation in terms of the three metrics considered.

**Table 12. Modeling Altruism Using Different Regression Models.**

| Model Type | Correlation | RMSE | MAE |
|---|---|---|---|
| **Demography Only** | 0.33 | 14.69 | 11.67 |
| ***Phoneotype* Only** | 0.75 | 10.26 | 7.99 |
| **Both** | 0.81 | 9.18 | 7.24 |

The demography based model obtained a correlation of 0.33 between the predicted and actual altruism values and RMSE and MAE scores of 14.69 and 11.67 respectively. The low - but significant - scores for the "demography only" model indicate that the demographic features can explain some (but not a lot) of variance in the altruism levels. Phone-based model performed much better with a correlation score of 0.75 (RMSE=10.26; MAE=7.99).

The combined model using *phoneotype* and demography features performed the best in terms of all three metrics and the predicted altruism was found to have 0.81 correlation with the actual altruism scores (RMSE=9.18; MAE=7.24). An MAE of 7.24 implies that the predictions are within ±7.24 of the absolute value of the altruism scores obtained by the survey (ground truth). Since the altruism scores obtained by the survey vary between 31 and 95 as shown in Table 10, ranges of ±7.24 can reflect a quite reasonable approximation.

Also, we see that the *phoneotype* model and "*phoneotype* + demography" (both) models yield considerably better models than the demography-based model. However, the demographic features were useful in increasing the correlation score for the *phoneotypic* model, thus suggesting that *phoneotypic* features and demography features are not merely proxies for each other, but rather add newer information when combined.

### 3.4.2. Building a Predictive Classification Model for Altruism Propensity

We aim to build and test a classification model capable of predicting altruistic propensities. The literature suggests various methods to cluster (group or categorize) such data, including standard median splits and extreme group analysis [129]. Standard median splits dichotomize continuous variables into two groups: “low” which is lower than the median value (50) of the data and “high” which is greater than the median as we have done in the previous chapter. However, such median splits have limitations in the sense they are often unable to capture the underlying dynamics of the observed phenomena because they are variable-oriented and not people-oriented [128]–[130].

Hence, in this chapter, we are going to use “unsupervised machine learning” to cluster the participants into naturally occurring groups based on their altruism scores. Specifically, we use k-Means++ [127] clustering algorithm to find the optimal clusters. The algorithm is initialized by choosing the first center randomly. Then, the succeeding centers are selected from the remaining points based on the squared distance from the closest center. We ran the algorithm ten times with 300 maximum iterations per each algorithm run. An important consideration for k-Means++ algorithm is the choice of the number of clusters (k) to be used by the algorithm. Literature suggests multiple methods including: “Silhouette scores”, “Bayesian Information Criteria (BIC)”, and “elbow method” for identifying the right number of clusters [135].

Here, we considered two different methods (Silhouette scores and BIC) for this process. Silhouette scores (higher score is better) compare the average distance to elements in the same cluster with the average distance to elements in other clusters [108]. We implemented this procedure in R 3.4.1 [103] and its package ClusterR 1.0.6 [136]. We found that the best k equals two as shown in Figure 3. This method generated two clusters, one of which contains the altruism scores 31 to 58 (N=36 participants), second of which contains the rest of the scores (60 to 95) (N=19 participants). For the ease of interpretation, we refer to these groups as “altruism group A” and “altruism group B” respectively.

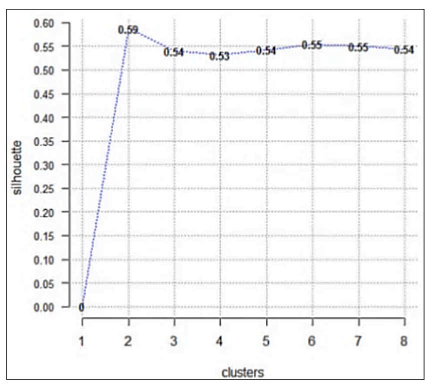


**Figure 3: Optimal Number of Clusters for k-Means++ (Silhouette Score).**

Using the BIC criteria (lower score is better) with the same ClusterR package to identify the optimal number of clusters for k-Means++, however, suggested the optimal number of clusters to be three. (Please refer to Figure 4). The first identified cluster contains the altruism scores from 31 to 45 (N=17 participants), second cluster contains the altruism scores 46 to 62 (N=24 participants), and the third cluster contains the rest of the scores (65 to 95) (N=14 participants).

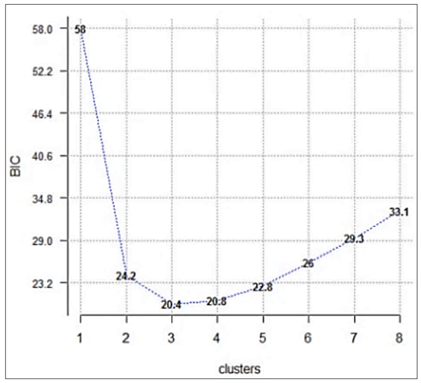


**Figure 4. Optimal Number of Clusters for k-Means++ (BIC).**

We used Orange3-3.6.0 [108], [137] to build the models which could automatically identify the altruism group (e.g. "altruism group A" or "altruism group B"), which an individual belongs to. We built 3 types of models based on k=2: one with the demographic features only, another one with the *phoneotypic* features only, and a third one with a combination of both types of features. We used information gain (reduction of entropy) [108] to rank the best subset of the (24 *phoneotype* + 6 demography) features described in the preceding section. For optimal feature subset selection, we first ranked all the features based on information gain [108]. Then, we considered models of up to ten features (a third of the available pool) wherein each model was the collection of top "j" features and j is in the range (1, 10). The optimal subset was the one with highest performance amongst the considered models. The resulting feature sets in each of the cases is shown in Table 13.

**Table 13. Features Selected for Various Prediction Models (k=2).**

| | |
|---|---|
| **Demography Only** | Income, School (Level of Education) |
| ***Phoneotype* Only** | Total Call Duration, Missed Call Percentage, Number of New Contacts (SMS) |
| **Both** | Total Call Duration, Income, Missed Call Percentage, Number of New Contacts (SMS), School |

The abovementioned features were used to test out several well-known machine learning algorithms for classification. Specifically, we used Naïve Bayes, Random Forest, CN2 Rule Induction, Logistic Regression, and kNN (k-Nearest Neighbors) with a leave-one-out cross validation method to balance between the learning opportunities and the generalizability of results from the data. Also, we used Zero-R (Constant or Majority) without any cross validation which simply classifies all the instances into the majority class as a baseline to facilitate interpreting the results. Table 14 offers a comparison of the results. It is worth noting that AUC stands for (Area Under the Receiver Operating Characteristic Curve) and CA means (Classification Accuracy) [108]. Moreover, F1 represents the harmonic mean between precision and recall [109]. A higher score (closer to 1) is better in each case and the Zero-R scores give a sense of the baseline expected performance.

**Table 14. Prediction Results for Altruism Levels Using Different Algorithms (k=2).**

| Method | Demography Only | | | *Phoneotype* Only | | | Both | | |
|---|---|---|---|---|---|---|---|---|---|
| | *AUC* | *CA* | *F1* | *AUC* | *CA* | *F1* | *AUC* | *CA* | *F1* |
| **NaïveBayes** | 0.594 | 0.655 | **0.486** | **0.756** | 0.727 | 0.595 | **0.798** | **0.782** | **0.667** |
| **RandomForest** | 0.594 | 0.673 | 0.471 | 0.664 | 0.636 | 0.375 | 0.679 | 0.691 | 0.452 |
| **CN2 Rule** | **0.656** | **0.709** | 0.429 | 0.508 | 0.618 | 0.276 | 0.587 | 0.582 | 0.343 |
| **LogRegression** | 0.558 | 0.582 | 0.303 | 0.586 | 0.618 | 0.160 | 0.598 | 0.600 | 0.154 |
| **kNN** | 0.459 | 0.655 | 0.000 | 0.719 | **0.782** | **0.667** | 0.719 | **0.782** | **0.667** |
| **Zero-R** | 0.500 | 0.655 | 0.518 | 0.500 | 0.655 | 0.518 | 0.500 | 0.655 | 0.518 |

Table 14 shows that the demography-based model returned the best classification accuracy (CA) of 70.9%, top AUC of 0.656, and F1 score of 0.486. The *phoneotype*-based model generated a better model than the demography model whose best accuracy is 78.2%, AUC is 0.756, and F1 is 0.667. The best results however were obtained by a combined model (demography + *phoneotype* data) that yielded an accuracy of 78.2%, AUC of 0.798, and F1 of 0.667.

We repeated a similar process of building three predictive models based on k=3. We found that the best results were obtained when we selected the following features in each of the models as shown in Table 15.

**Table 15. Features Selected for Various Prediction Models (k=3).**

| | |
|---|---|
| **Demography Only** | Income, School (Level of Education), Age |
| ***Phoneotype* Only** | Missed Call Percentage, Weak Ties (SMS), Weekday Weekend SMS Ratio, Number of New Contacts (SMS), Weak Ties (Call) |
| **Both** | Income, Missed Call Percentage, Weak Ties (SMS), School, Age, Weekday Weekend SMS Ratio, Number of New Contacts (SMS), Weak Ties (Call) |

Table 16 shows that the demography-based model returned the greatest accuracy (CA) of 45.5%, top AUC of 0.577, and F1 score of 0.517. The *phoneotype*-based model generated a better model than the demography model whose best accuracy is 56.4%, AUC is 0.777, and F1 is 0.717. The best results were obtained by a combined model (demography + *phoneotype* data) that yielded an accuracy of 65.5%, AUC of 0.816, and F1 of 0.760.

**Table 16. Prediction Results for Altruism Levels Using Different Algorithms (k=3).**

| **Method** | **Demography Only** | | | ***Phoneotype* Only** | | | **Both** | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***AUC*** | ***CA*** | ***F1*** | ***AUC*** | ***CA*** | ***F1*** | ***AUC*** | ***CA*** | ***F1*** |
| **NaïveBayes** | 0.535 | 0.273 | 0.279 | **0.777** | **0.564** | **0.717** | **0.816** | **0.655** | **0.760** |
| **RandomForest** | 0.519 | 0.382 | 0.431 | 0.569 | 0.455 | 0.538 | 0.558 | 0.364 | 0.464 |
| **CN2 Rule** | **0.577** | **0.455** | 0.435 | 0.624 | 0.436 | 0.576 | 0.639 | 0.473 | 0.571 |
| **LogRegression** | 0.526 | 0.436 | **0.517** | 0.414 | 0.255 | 0.393 | 0.528 | 0.491 | 0.566 |
| **kNN** | 0.507 | 0.418 | 0.412 | 0.593 | 0.418 | 0.500 | 0.606 | 0.436 | 0.528 |
| **Zero-R** | 0.500 | 0.436 | 0.265 | 0.500 | 0.436 | 0.265 | 0.500 | 0.436 | 0.265 |

Note that these results in terms of accuracy are lower for the three-way classification problem compared to the two-way classification problem. However, a three-way classification problem is in general a harder problem than two-way classification, and the much lower baseline (Zero-R) scores may help interpret the performance gain obtained by the phone-based models.

From the aforementioned results, we can clearly observe that *phoneotypic* features considerably outperform demography-based ones in prediction's accuracy corroborating the findings from the regression analysis. Further, similar to the regression analysis, the *phoneotypic* (i.e., phone-based) behavioral features were not merely a replacement for demographic features, as the combined models yielded higher performance as compared to the individual models.

It is also clear that the *phoneotypic* model outperformed the baseline Zero-R model. In the case of k=2, the *phoneotypic* model performed 59.6% better than the baseline model in terms of AUC, 19.4% better in terms of accuracy, and 28.8% better in terms of F1 score. In the case of k=3, the *phoneotypic* model performed 63.2% better than the baseline model in terms of AUC, 50.2% better in terms of accuracy, and 186.8% better in terms of F1 score.

Hence, we note that a phone-features based model beats baseline majority classification and also goes beyond static demographic descriptors (e.g. age, gender, education, income) for predicting altruism propensities. This underscores the potential for using phone-based (or *phoneotypic*) features to build automatic classifiers for individual altruism propensities. These findings provide clear evidence of interconnections between the mobile features and altruistic propensities and also motivating further work in this direction. In effect, these results pave way for personal big data to expand the understanding of the associations between socio-mobile behavioral data and altruism propensities.

**3.4.3. Behavioral Features Associated with Altruism Propensity**

Besides identifying automated methods for identifying an individual's altruism propensity levels, one of the goals of this work is to understand the socio-mobile behavior of individuals with different propensities to be altruistic. Thus, we undertook a *post-hoc* Pearson's correlation analysis using IBM SPSS 24 between the altruism scores obtained from the survey and the *phoneotypic* features. Note that the correlation analysis undertaken here is *posthoc* and intended to help interpret the observed predictions, as opposed to being prescriptive in its own right.

In the interest of space, we only report the correlations that were found to be (at least marginally i.e. $p<0.10$) significant in Table 17. We note that people who have high altruism propensity tend to be more socially active, yet have different usage patterns for different communication modalities.

First, individuals with higher altruism propensity tend to call more often ($r$= +0.356). This can be understood as altruism propensity being associated with healthy social relationships and higher call activity captures such behavior [8]. We see a similar trend in terms of new call contacts ($r$= +0.305). This underscores the importance of constantly renewing and broadening one's social contacts and its associations with altruism. From a methodological perspective the dynamics of social contacts captured by this feature underscore the value of temporal features, which cannot be captured in one-time lab studies studying the phenomena of altruism.

Next, we notice that individuals with higher altruism propensity show a marked preference for engaging in phone calls with their "strong ties" as opposed to spending it equitably with all contacts (*r*= +0.252). Conversely, they spend less time on calls with their "weak ties" (*r*= -0.232). However, the patterns of SMS communication seem to be quite different from phone call-based communication. SMS interactions with "weak ties" were found to be positively associated (*r*= +0.260) with altruism. From a methodological perspective, these results suggest the value of different modalities of data to triangulate and predict human traits. From a conceptual perspective, these observations corroborate previous studies which suggest that altruism is a pro-social trait associated with higher social capital including both its bridging and bonding variants [134], [92]. The correlation scores for the significant features in Table 17 were also found significant in LASSO regression as explained earlier.

**Table 17. Correlation between Phone-based Features and Altruism Propensity (** 0.01, * 0.05, ° 0.10).**

| ***Phoneotypic* Feature** | **Pearson's Correlation** | **Significance (p-value)** |
|---|---|---|
| **Social Activity Level (Call)** | 0.356** | 0.008 |
| **Number of New Contacts (Call)** | 0.305* | 0.024 |
| **Strong Ties (Call)** | 0.252° | 0.064 |
| **Weak Ties (Call)** | -0.232° | 0.089 |
| **Weak Ties (SMS)** | 0.260° | 0.055 |

### 3.5. Discussion

The three forms of analysis implemented in this work (regression analysis, classification, and Pearson's correlation analysis) suggest that machine learning, and data analytics approaches generally, can be utilized to infer individual altruism propensity based on phone metadata to a substantial extent. The regression analysis can estimate the individual altruism propensities with high correlation (0.75) and within a margin of ±7.99 over a range of 31 to 95. Accompanying phone features with demographic data, where available, could yield even better performance. For instance, the classification analysis yielded up to about 80% accuracy (AUC=0.782; F1=0.667) based on such combination of *phoneotype* and demography features.

Given the small sample size, we focus here on exploring general patterns and trends over the three analysis techniques (regression, classification, and correlation). We can observe a consistency in the results across the three analysis types as well as the two variants (k=2 and k=3) for classification, suggesting that socio-mobile signals as observed via a phone (*phoneotype*) could indeed be used to infer altruism propensity of an individual. The results contribute to the growing literature on using "personal big data" to characterize multiple traits of human beings [23], [54], [68], [69], [71]. At the same time, they motivate further work to study altruism propensity using socio-mobile behavioral data.

### 3.5.1. Privacy of User Data and Ethical Considerations

To insure and maintain the privacy of the participants, we followed the best practices in the human subjects' research that require hashing and anonymizing all data before analysis. Also, no one from the research team under any circumstance had an access to private data like the exact phone number of a participant or the content of the calls or SMS messages. The Android app collecting the data requires lesser permissions than many of the popular apps available at Google Play Store (e.g. WhatsApp, Instagram).

We also note the moral and ethical considerations in giving a person a score based on their propensity to be altruistic. History repeats itself; similar reservations have been raised up about the conventional paper survey approaches with a similar objective, and likewise automatic systems which use social media and phone data to give health, well-being, or similar scores to people [138]. Rather than waiting for perfect privacy and ethics guidelines to emerge around these topics, we posit that studies like ours can help broaden the understanding around the prospects of using "personal big data" to create personalized sociological profiles of individuals and inform the discussion in the research community around them [139].

### 3.5.2. Limitations

The work in this chapter has three limitations: (1) homogeneity of the sample (participants were mostly undergraduate students from the same institution), (2) small sample size (55) while having large number (24) of potentially collinear features in regression analysis, and (3) Inability to establish causality. Bearing in mind these limitations, we will be cautious in generalizing the findings obtained until they are verified them at scale over representative sample populations. To overcome these limitations, we used LASSO regression which deals with such situations of having relatively more number of features for a sample trying to minimize overfitting by penalizing the use of too many features [102]. Furthermore, we plan to repeat this work in the future considering a larger and more diverse sample.

Despite these limitations, to the best of our knowledge, this is the first line of work to analyze the links between altruism levels and phone-based socio-mobile behavior (*phoneotype*). The obtained results in this first of its kind effort are thus encouraging, and have demonstrated the potential of personal big data for predicting altruism levels of individuals.

**3.5.3. Implications**

The results open the doors to a methodology that, with refinements and validation, could be used at scale. Smart phones are now actively used by more than 1.4 billion users, and hence the proposed method could potentially be applied to estimate the altruistic levels for billions of individuals.

In future, this work could also have multiple implications for social scientists, economists, mobile phone service providers, and policy designers. For example, the suggested methodology could help social scientists study altruism at scale in the society. Besides identifying connections between spatial and temporal behavior and altruism, a scalable methodology to study altruism could allow for asking questions regarding the spread of altruism in networks of billions of individuals, which are simply not possible with current survey or lab-based methods.

Similarly, as mobile phones are increasingly used, both, as user end-points and as mediators of technology, modeling a person's altruistic propensities automatically could be helpful in supporting various socio-technical applications under the Internet of People vision [18]. Such an Internet-of-People vision explicitly requires the creation of a "sociological profile" [18] for the participants and the proposed method for inferring altruistic propensities could be used, for example, to identify a person's default preferences in peer-to-peer networking, file sharing, or human-computation based tasks. Further, with multiple bots negotiating services and conditions for users in the emerging social Internet of Things scenarios, having such a sociological profile could be useful to suggest default settings in multiple scenarios, from something as simple as setting the right room temperature in shared workspaces to the default “tipping” amount in dinner payments.

# Chapter 4

# Modeling Interpersonal Trust Using Deep Learning

## 4.1. Introduction

Interpersonal trust is defined as “a willingness to accept vulnerability or risk based on expectations regarding another person’s behavior” [10]. It facilitates various socio-technical systems with many implications affecting personal and societal well-being. The emerging growth of social networks, ubiquitous sensors, and social internet of things, necessitate understanding and modeling of people’s interpersonal trust as they interact with one another for undertaking tasks ranging from mobile commerce to shared economy transactions. For instance, a person may want to obtain recommendations only from somebody they trust, rent out homes only to somebody they trust, or seek child care services only from somebody they trust. Each of these aspects is already being mediated by mobile phone apps (e.g. Amazon, Airbnb, TaskRabbit, UrbanSitter) and as the trend is only likely to increase with the emerging internet of things, modeling interpersonal trust is a critical problem for ubiquitous computing research.

Recently, there has been a tremendous interest in utilizing personal ubiquitous data to automatically infer different attributes about a person or their interpersonal relationships [23], [84], [140]. Multiple recent efforts in particular have focused on defining deep learning approaches for inferring aspects of an individual using phone and ubiquitous data [59], [60], [141]. We see great potential in this line of work and propose to extend it based on some well-known insights from the social science literature. These insights include the concepts of homophily i.e. *birds of a feather flock together,* and *word of mouth* [142]. Homophily suggests that an individual's traits and attributes can be inferred better if we also observe the attributes of their neighbors and multiple studies have shown that the *word of mouth* i.e. communication with close neighbors is an important determinant of whom one trusts [142]–[144].

While current deep learning architectures are typically well-designed to handle low level notions of neighborhood within an entity of interest (e.g. neighboring pixels *within* an image or the next Bluetooth reading within a stream), they typically do not consider the inter-entity notions of neighborhood. While this may be less of an issue when dealing with intra-entity problems e.g. labeling objects within an image, this becomes an important limitation in tackling problems related to human relationships as they are almost always affected by neighboring relationships within the same network. **To the best of our knowledge, there has been no previous work in ubiquitous computing literature that utilizes the data coming from neighboring relationships in a deep learning architecture to better infer the properties for a target relationship**. While the proposed approach will have implications for multiple social inference problems, here we ground and test it in the context of interpersonal trust prediction where based on the *word of mouth* rationale, one's trust in a target entity could be strongly influenced by the *relationship between one's neighbors and the target entity*.

We implement and test the proposed architecture on the MIT Friends and Family dataset [71], which contains interpersonal trust scores as well as Call, SMS, and Bluetooth based social interactions between a community of 130 users for a period of one year. We apply multiple shallow and deep learning approaches to use phone based behavioral data (Call, SMS, and Bluetooth) to automatically infer the interpersonal trust scores between any two members of the community. Based on the analysis, the trends indicate that:

(a) Adding information about neighbors yields better performance at inferring interpersonal trust in both shallow and deep learning approaches.

(b) Deep learning approaches perform better at inferring interpersonal trust than comparable shallow approaches.

(c) The proposed deep learning architecture, which is aware of the interaction effects between neighbors (NADAL) yields higher performance than a baseline feature concatenation based deep learning (FC-DNN) approach for combining information coming from neighbors.

Building upon the use of mobile phone metadata, the results pave way for understanding interpersonal trust at a societal scale and have implications for numerous applications in the emerging social internet of things (e.g. Uber, TaskRabbit, Meetup) as well as almost any human task that involves two people to cooperate with each other.

**4.2. MIT Friends and Family Dataset**

MIT Friends and Family dataset was part of a yearlong study utilizing “Funf in a box” framework [71] and surveys to collect data about the lives of 130 individuals (about 64 families) living in a families only housing on campus of a major North American University. The Funf platform is capable of collecting various types of data though here we focus on the Call, SMS, and Bluetooth logs as well as the trust surveys that determine trust ties between the subjects [30].

To accommodate the various definitions of trust occurring in three important hypothetical but pertaining to daily life scenarios (well-being, money, and kinship), the participants were asked the following three questions [30]:

(1) “Would you ask person X for help in sickness?”

(2) “Would you ask person X for a hundred-dollar loan?”

(3) “Would you ask person X for babysitting?”

We focus in studying the third question only in this study due to missing data in the first two questions in the version available to us. To capture several aspects of human relations in the dataset, Bluetooth (BT), Calls, and SMS logs were collected. Explicitly, using call logs facilitates understanding the synchronous interaction between two individuals despite their distance. Also, using SMS logs enables understanding the non-synchronous interaction between two individuals regardless of their distance. Bluetooth logs facilitate understanding the spatial patterns of the participants in face to face interactions where the logs get updated every five minutes based on scanning for adjacent Android phones [30].

We focus only on the interactions which take place within the community (e.g. disregard external calls). The participants collectively made a total of 474,340 BT scans (proxies for face to face encounters) with an average of about 4351.74 and a median of 3864 per participant, made 58,554 calls with an average of around 476.05 and a median of 407 calls per participant, and exchanged 17,369 SMS messages with an average of 231.59 approximately and a median of 88 per participant during the period of the study. Table 18 gives a summary of the total, mean, and median for calls, SMS, and BT scans.

**Table 18. Summary of Calls, SMS, and BT.**

| Feature | Total | Mean | Median |
|---|---|---|---|
| **BT** | 474,340 | 4351.74 | 3864 |
| **Call** | 58,554 | 476.05 | 407 |
| **SMS** | 17,369 | 231.59 | 88 |

## 4.3. Mobile Phone Data Features

Trust and socio-mobile behavior have been connected in previous research both conceptually and empirically. For instance, an individual's propensity to trust others has been connected conceptually with social capital and empirically with phone data [145]. Similarly, interpersonal trust has been connected conceptually with the *strength of ties* and empirically with phone data [30], [146]. Interpersonal trust, as reported by person A, is often a function of the *trust propensity* of person A, as well as the *interactions* between person A and the target entity B. Hence, we consider features for both the node properties of A as well as the edge properties for (A→B) to model the interpersonal trust between A and B.

To come up with such a representation using social-mobile data, we surveyed the related literature which focuses on connecting phone behavior with social outcomes (e.g., [30], [55]). For example, social capital as a concept is connected with both phone use behavior [54] and trust [51]. Social capital often comes in two variants: bridging and bonding [45]. Hence, we link the concepts of weak and strong ties to bridging and bonding social capital to infer interpersonal trust [23], [93], [94]. We use Call, SMS, and Bluetooth logs to represent the features that carry "social traits" concepts for mobility and interpersonal trust and their interconnections [37], [54], [147]. Based on the BT, Call, and SMS metadata collected from the app, we define the following set of phone-based features (N=23) as presented in Figure 5 and Table 19.

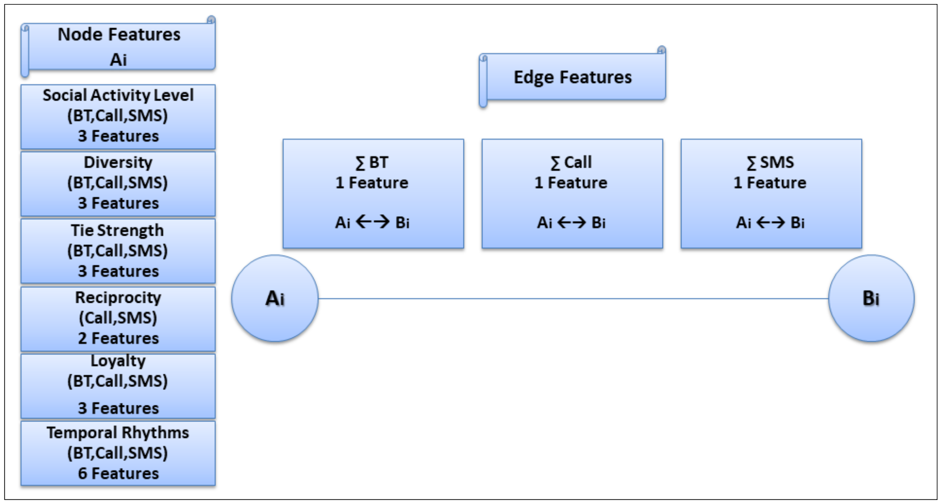


**Figure 5. Summary of Phone-based Features in a Network Representation.**

**Table 19. Summary of *Phoneotypic* (Phone-based) Features Defined in this Work.**

| Feature | Definition |
|---|---|
| **NODE: Social Activity Level 3 Features** | $\text{Social Activity (BT, Call, SMS)} = \sum Activity$ |
| **Diversity 3 Features** | $\mathrm{D_i} = -\sum_j p_{ij} \log_b p_{ij}$ |
| **Tie Strength 3 Features** | $\text{Strong/Weak Tie Ratio (SWTR)} = \dfrac{\left(\frac{\sum Communication\ with\ Highest\ 1/3\ Contacts}{\sum All\ Communication} X\ 100\right)}{\left(\frac{\sum Communication\ with\ Lowest\ 1/3\ Contacts}{\sum All\ Communication} X\ 100\right)}$ |
| **Reciprocity 2 Features** | $\text{In Out Ratio (Call, SMS): IOR} = \dfrac{Incoming\ Communication\ Count}{Outgoing\ Communication\ Count}$ |
| **Loyalty 3 Features** | $\text{Loyalty} = \dfrac{\sum Time\ Spent\ with\ Top\ Three\ Contacts}{\sum Time\ Spent\ with\ All\ Contacts} X\ 100$ |
| **Temporal Rhythms 6 Features** | $\text{Diurnal Activity Ratio (BT, Call, SMS) DAR} = \dfrac{\sum Activity\ (8AM\ to\ 5PM)}{\sum Activity\ (5PM\ to\ 8AM)}$<br>$\text{Weekday/Weekend Activity Ratio (BT, Call, SMS) WWAR} = \dfrac{\sum Activity\ (Weekdays)}{\sum Activity\ (Weekends)}$ |
| **EDGE: Social Activity Level 3 Features** | $\text{Social Activity (BT, Call, SMS)} = \sum Activity$ |

Note that while many deep learning approaches do not utilize "hand crafted" features, there remain multiple scholars who have argued that theory driven (or hand-crafted) features are useful even when using deep learning architecture [148]–[150]. While availing of the sophisticated non-linear interactions between features using the neural networks, such approaches still allow the system designers to have a better understanding of the rationale for their models. Further, such features allow for a more interpretable comparison between shallow and deep learning approach results, and work well in scenarios where the available number of instances is not exceptionally large [151]. In the current scenario, where there is only interpersonal relationship score per edge even though they may interact for over a year we have opted to use hand crafted features at the input layer. This also allows for a comparison across deep and shallow learning strategies for using neighbor's sensor data for inferring interpersonal trust – neither of which has been reported in the past literature.

### 4.3.1. Node Properties

#### 4.3.1.1. Social Activity Level

Social Activity level represents the activity of a user as obtained through counting exchanged phone calls, SMS messages, and Bluetooth scans. A higher count of social activity level suggests an active user [23]. Various studies have connected individual social activity with social capital and/or trust. High social activity has also been connected with reducing relational uncertainty and as a means of establishing trust in interpersonal relationships [30], [88]. We assume that a person who makes or receives (I/O) numerous calls may have more social life and this may be associated with interpersonal trust [152]. Also, we consider the level of social activity at the edge (vertex) level between the two users whose interpersonal trust score is inferred. Thus, we consider the following features:

$$\text{Social Activity (BT, Call, SMS)} = \sum Activity$$

#### 4.3.1.2. Diversity

In the previous set of features, we quantify the total number of calls, SMS messages, and Bluetooth scans. Here, we also determine the diversity (measured as Shannon Entropy) for each one of them, as such a diversity metric has been reported to be associated with various personal well-being outcomes and personality traits [69], [90].

$$\mathrm{D_i} = -\sum_j p_{ij} \, log_b \, p_{ij}$$

Where $p_{ij}$ is the percentage of social events involving individual 'i' and contact 'j', and 'b' is the total number of such contacts.

#### 4.3.1.3. Tie Strength

Previous studies have related strength of ties and trust [92]. Such literature underscores the value of maintaining relationships with both strong and weak ties, and each may yield different types of social capital, and presumably, over periods of time, interpersonal trust. Following Williams [93], we connect the concepts of 'bonding' and 'bridging' social capital to those of 'strong' and 'weak' ties as proposed by Granovetter and other researchers [94]–[96]. We conjecture that the relative spread (or concentration) of communication with strong (respectively weak) ties may be a predictor of one's interpersonal trust to others. It is anticipated that a person would devote at least 33% of their time with their top-third most frequent contacts (proxy for strong ties) [23]. Nonetheless, a high score like 90% may indicate an individual's preference to intentionally engage more with strong ties rather than distributing the communication effort more equally amongst all ties. Hence, we define the following features:

$$\text{Strong/Weak Tie Ratio (SWTR)} = \frac{\left(\frac{\sum Communication\ with\ Highest\ 1/3\ Contacts}{\sum All\ Communication} X\ 100\right)}{\left(\frac{\sum Communication\ with\ Lowest\ 1/3\ Contacts}{\sum All\ Communication} X\ 100\right)}$$

**4.3.1.4. Reciprocity**

The ease with which communication is conducted is an important property of an individual's social behavior. We anticipate approachability of individuals to be associated with social capital levels [54]. Such social capital levels have been associated with trust [14]. Hence, we compute the ratio of incoming to outgoing calls and SMS text messages.

$$\text{In Out Ratio (Call, SMS): IOR} = \frac{Incoming\ Communication\ Count}{Outgoing\ Communication\ Count}$$

**4.3.1.5. Loyalty**

Loyalty means how frequently participants engage with their favorite people in terms of Calls, SMS messages, and Bluetooth scans. Past research has connected this loyalty feature with individual well-being and propensity to trust [98], [145]. Precisely, we calculate the percentage of time spent with their top three frequented communication (BT, Call, SMS) out of all communication.

$$\text{Loyalty} = \frac{\sum Time\ Spent\ with\ Top\ Three\ Contacts}{\sum Time\ Spent\ with\ All\ Contacts} X\ 100$$

**4.3.1.6. Temporal Rhythms**

Prior literature has connected circadian cycles, Dark Triad (i.e., narcissism, machiavellianism, and psychopathy) and trust [81], [82]. The classification of different individual's chronotype - the tendency for the individual to sleep at a particular time during a day-and/or-night period (24-hour) - has been connected with cheating and machiavellianism [83]. The daily business hours in USA are 8 AM to 5 PM; hence, to quantify daily patterns of activity and the differences between different phases of the day, we define the following features:

$$\text{Diurnal Activity Ratio (BT, Call, SMS) DAR} = \frac{\sum Activity\ (8AM\ to\ 5PM)}{\sum Activity\ (5PM\ to\ 8AM)}$$

We added another layer of characterization for the aforementioned two states of the daily activity ratio to get more insights out of these circadian rhythms by quantifying the weekdays (Monday to Friday) to weekends (Saturday and Sunday) communication (BT, Call, SMS) ratio.

$$\text{Weekday/Weekend Activity Ratio (BT, Call, SMS) WWAR} = \frac{\sum Activity\ (Weekdays)}{\sum Activity\ (Weekends)}$$

### 4.3.2. Edge Properties

Past research has connected the number of interactions between users conceptually with the strength of ties [22], [95] and this feature has empirically been found to be predictive of interpersonal trust [30]. Hence, we consider Social Activity Level based on the three modalities (BT, Call, SMS) as the features to characterize the edges in the network as seen in Figure 5.

## 4.4. Proposed Method

### 4.4.1. Dealing with Class Imbalance in the Dataset

To clean our dataset, we have removed all instances without any logs for BT scans, Calls, and SMS messages altogether resulting in a total of 13,164 instances. 98.75% of these instances have a low interpersonal trust (zero), whereas the rest of instances (1.25%) have a high interpersonal trust (one) as shown in Figure 6. Most common way of dealing with class imbalance in datasets is to (artificially) balance the training set to allow for better learning opportunities, before the learnt model is tested out on the imbalanced test as is expected in the real world. The most common ways for balancing the training dataset are: (1) over-sampling, (2) under-sampling, and (3) combination of over and under-sampling [153]. Here, we split the dataset into two (train/test) subsets (70% - 30%) respectively and chose SMOTE+Tomek links [154] to balance the training data. In this technique, SMOTE [155] is used first which generates a new minority class instances. These minority class instances are based on a projection in the hyperspace and are not direct copies of the existing instances. Then, Tomek's links method is used to under-sample the dataset whose main motivation is not only to balance the training data, but also to remove noisy examples lying on the wrong side of the decision border [154]. We use the implementation as described in [153], which is inspired by [156].

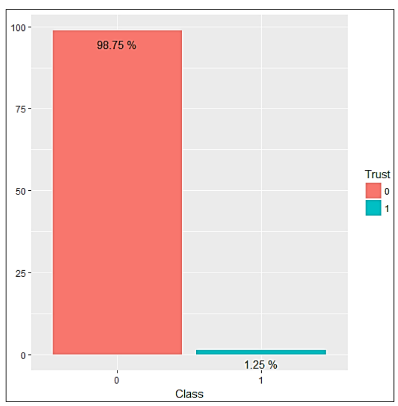


**Figure 6. Class Imbalance of the Interpersonal Trust Scores.**

### 4.4.2. Identifying Appropriate Neighbors for Better Interpersonal Trust Modeling

In this work, we would like to study the novel idea of determining the impact of neighbors in enhancing the performance of shallow and deep learning algorithms in inferring interpersonal trust. Figure 7 displays the network (indirect graph) of the users of the dataset based on Bluetooth scans between them during the course of the study. Following the results in [30], we consider face-to-face interactions to be the most important determinants for considered trusted relationships and hence identify the two nearest neighbors in terms of the frequency of face to face (Bluetooth) interactions. The construction of the network was done in R 3.5.1 [103] and its package iGraph 1.0.1 [157] inspired by [158].

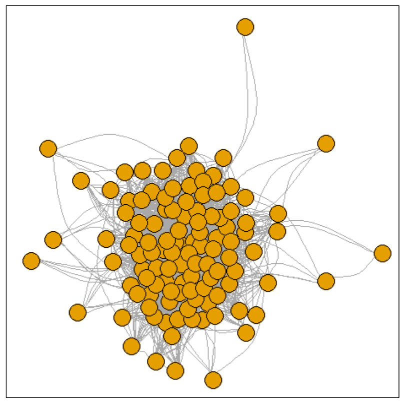

**Figure 7. Network (Graph) of the Dataset Based on BT Scans.**

The underlying intuition of using neighbor's information can be applied to any arbitrary number of neighbors. However, including all the neighbors would quickly become exorbitant in terms of data size and the effects of additional user data are unlikely to be useful after a threshold. Given the significance of triads as an important building block in social network literature [159], [160] and "triangulation" in signal processing literature, we focus on the use of additional data from two neighbors in this work.

To study the impact of adding the two similar neighbors based on number of BT scans in the prediction of one's interpersonal trust, we have created the following additional features as presented in Figure 8.

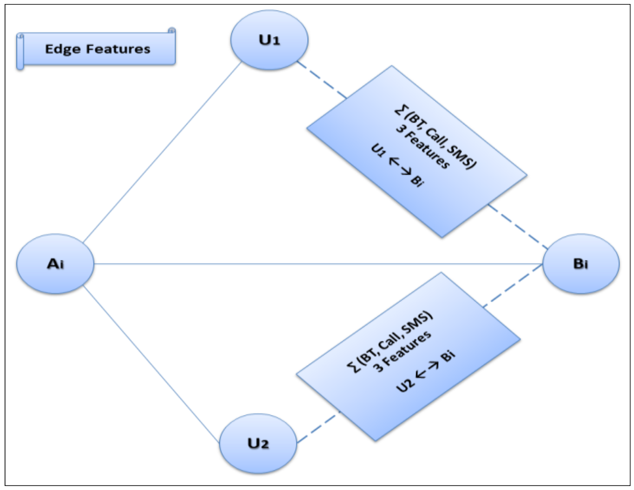


**Figure 8. Additional Phone-based Features in a Network Representation after Adding Two Neighbors.**

### 4.4.3. NADAL: A Neighbor Aware Deep Learning Architecture

In this work, we build upon a recent work in Ubicomp literature by Radu et al. [161], which defines a novel deep learning approach to utilize multimodal sensor data for human activity recognition. An important insight from their work was the idea to avoid both the extremes of fusion techniques i.e. early fusion (feature concatenation) and late fusion (decisions derived separately from single modalities are combined in the final layer). Instead, they argue a case to allow for two types of hidden layers: hidden layers related to a specific sensor type and hidden layers that capture unified concepts across sensor types. In their construction, separate architectures are built for each modality to first learn sensor-specific information before their generated concepts are unified through representations that bridge across all the sensors (i.e., shared modality representations).

We consider the data coming from the neighbors to be an additional "channel" or modality of information regarding the phenomena of interest. In that sense, our work follows that of Radu et al. [161]. However, the "channels" in our setup are quite different from those in Radu et al. [161]. While in their context, different channels were observing the *same* activities via different modalities, in our case the additional channels provide contextual information regarding *different* activities, which nevertheless could indirectly influence the prediction task at hand.

Specifically, we consider the interpersonal trust between user $A_i$ and a target $B_i$ (see Figure 8) to be a function of the behavioral features which characterize the edge $\{A_i \rightarrow B_i\}$ (e.g. the number of phone calls between them) as well as the node $A_i$ (e.g. number of overall phone calls made by $A_i$). While the node properties give a clue to the personality or the traits of $A_i$, the edge properties characterize the relationship between $A_i$ and $B_i$.

Now, let us also consider two neighbors for $A_i$: $U_1$ and $U_2$. We posit that the properties of the edges connecting these users with $B_i$ i.e. $\{U_1 \rightarrow B_i; U_2 \rightarrow B_i\}$ could provide additional context on the relationship $\{A_i \rightarrow B_i\}$ and thus be useful to predict the interpersonal trust between them. However, we do not expect the node properties (e.g. personality or trust propensity) of $U_1$ and $U_2$ to have a significant influence on the relationship between $A_i$ and $B_i$.

Hence, taking inspiration from Radu et al. [161] but also considering the different application context here, we define a novel architecture as shown in Figure 9. This architecture builds upon feedforward neural networks and contains separate architectural branches for user $A_i$'s node and edge features as well as $U_1$'s and $U_2$'s edge features without any inter-branch connections between layers until later unifying cross channel layers connect the node and edge features for the $A_i$, and the three types of edges respectively. It allows for the node properties of user $A_i$ to go through a number of layers of neural networks to allow for different features and the interactions among them to become part of the model. Same thing happens to the other "channels" of information i.e. the edge properties of $A_i \rightarrow B_i$, $U_1 \rightarrow B_i$, and $U_2 \rightarrow B_i$. Each of these properties goes through a number of layers of neural networks without any interaction across channels. Next, to learn the (potentially non-linear) interaction effects between the $A_i$'s node and edge parameters, the corresponding layers are merged and the resulting layer passes through multiple layers of networks to allow for learning of the appropriate parameters. Similarly, there could be interaction effects between the edge-based features for $A_i \rightarrow B_i$, $U_1 \rightarrow B_i$, and $U_2 \rightarrow B_i$, which can be learnt by combining the corresponding layers and letting them pass through two layers of neural networks.

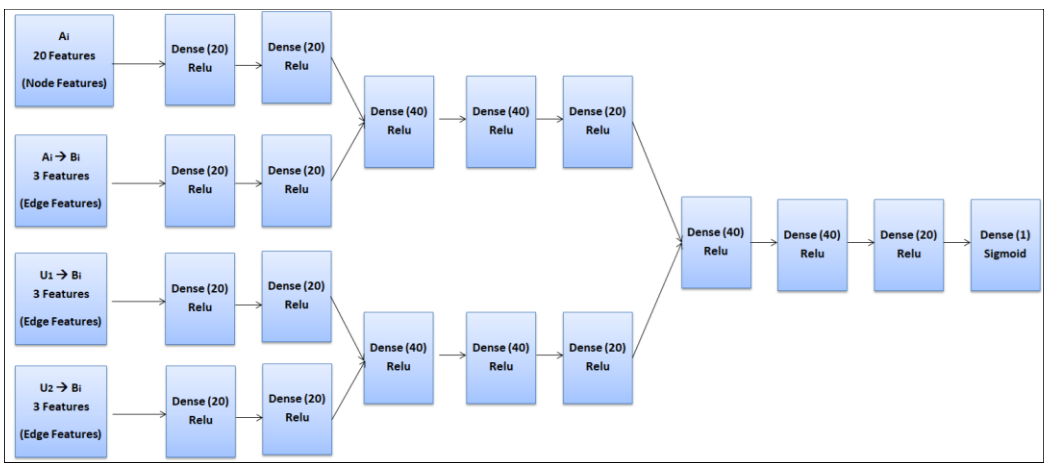


**Figure 9. NADAL Architecture Schematic.**

## 4.5. Results

We used Scikit-learn [114] and Keras [162] running in Google Colab Notebooks to build various models capable of automatically inferring the interpersonal trust of a user classified into two classes: "Low Interpersonal Trust (0)" or "High Interpersonal Trust (1)". We split all datasets into 70% training dataset and 30% test set. We analyze the results with and without sampling, as well as with and without considering the neighboring edges. Lastly, we consider both shallow and deep learning methods, namely: Random Forest, standard feature concatenation based deep neural network (FC-DNN), and our proposed NADAL architecture.

(1) **Sampling Technique: As-Is vs. SMOTE+Tomek Resampling**

As mentioned in Section 4.4.1., we try to counter the problem of class imbalance by creating more balanced training datasets using SMOTE+Tomek resampling. To quantify the performance difference based on the re-sampling, we run two versions for each experiment – one with and one without the resampling.

(2) **Neighbor Awareness: Individual Path (Non-Neighbor-Aware) vs. Neighbor-Aware**

We consider the performance of the models if they only utilize the individual node and its edge connecting the target node features vs. utilizing the edge data from two of the closest neighbors. While all the individual path approaches had access to 23 features (20 node features + 3 edge features) the Neighbor-aware approaches had access to 29 (20 node features + 3 * 3 edge features). While the difference in number of features made little impact on the architectures for Random Forest and FC-DNN, the NADAL architecture was adapted to consider only the layers that lie in the path of the abovementioned 23 features for the computation.

(3) **Machine Learning Approach: Shallow Learning (Random Forest) vs. Deep Learning (FC-DNN and NADAL)**

We consider three types of machine learning approaches. First is Random Forest, which is a frequently used shallow learning technique. Next is the baseline deep learning approach, which builds upon feature concatenation in the first layer (FC-DNN) and lastly the NADAL approach, which has been custom designed to capture the interactions between edges of neighboring nodes. For FC-DNN, we passed all the features through a multilayer perceptron (23/40/40/20/40/40/1) all activated by Relu except the output layer that was activated by Sigmoid with a 16 batch size and a 50 epochs as presented in Figure 10. For NADAL, the features were passed through different layers as shown in Figure 9, where all layers are activated by Relu except the output layer which was activated by Sigmoid with a 16 batch size and a 50 epochs.

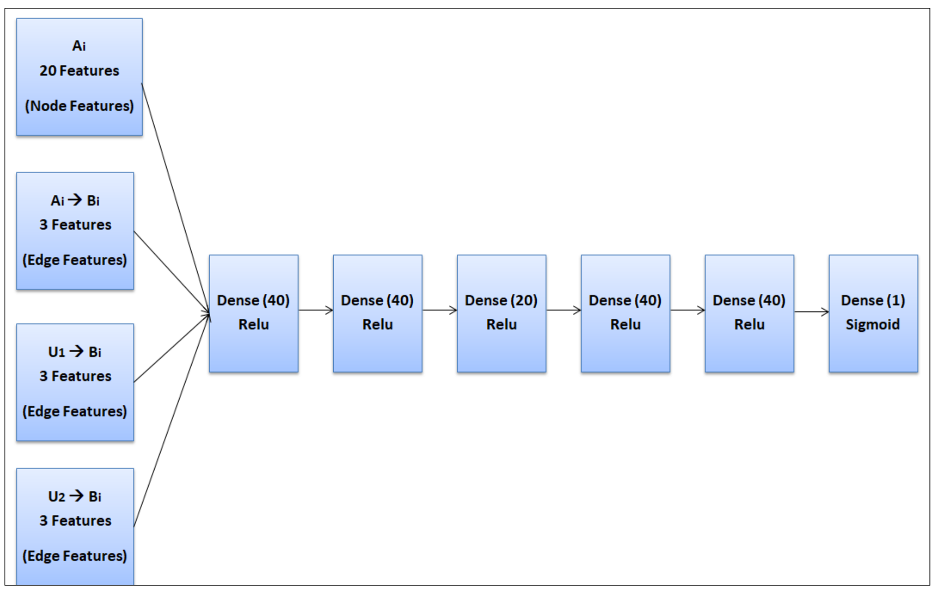


**Figure 10. Standard Feature Concatenation Based Deep Neural Network (FC-DNN) Architecture Schematic.**

Table 20 compares all datasets and the number of features of each one of them. Note that while the training data (70%) is balanced between the two classes by creating artificial samples (SMOTE+Tomek) the testing (30%) is done on the imbalanced dataset as consistent with the real world scenario where such an algorithm is likely to be applied.

**Table 20. Summary of Datasets Considered in This Study.**

| Dataset | Neighbor Awareness | Features | Instances | Class 0 | Class 1 |
|---|---|---|---|---|---|
| **ORIGINAL** | **Only Main Edge (100%)** | 23 | 13,164 | 12,988 | 176 |
| **ORIGINAL** | **Main Edge + Two_Neighbors (100%)** | 29 | 13,164 | 12,988 | 176 |
| **TRAINING SET: AS-IS** | **Only Main Edge (70%)** | 23 | 9214 | 9103 | 111 |
| **TRAINING SET: AS-IS** | **Main Edge + Two_Neighbors (70%)** | 29 | 9214 | 9103 | 111 |
| **TRAINING SET: After Resampling** | **Only Main Edge (70%)** | 23 | 18,170 | 9085 | 9085 |
| **TRAINING SET: After Resampling** | **Main Edge + Two_Neighbors (70%)** | 29 | 18,168 | 9084 | 9084 |
| **TEST SET** | **Only Main Edge (30%)** | 23 | 3949 | 3895 | 54 |
| **TEST SET** | **Main Edge + Two_Neighbors (30%)** | 29 | 3949 | 3895 | 54 |

Table 21 shows the average results of running experiments with each of the abovementioned settings 10 times. It is worth noting that AUCROC stands for (Area Under the Receiver Operating Characteristic Curve) and Acc stands for (Accuracy). While a higher score is better for each of these metrics generally, multiple researchers have suggested against using Classification Accuracy to interpret results in highly imbalanced datasets [109], [163]. For instance, a simple baseline (Majority Zero-R) algorithm which classifies all ties as “not trusted” will achieve an accuracy of 98.75%. However, such an algorithm would be useless in practice. Hence, we use AUCROC, which tries to balance the performance for the majority and the minority class as the primary metric to compare algorithms.

The results summarized in Table 21 show the following trends:

The use of SMOTE+Tomek re-sampling technique allowed for the algorithms to achieve better performance across all deep learning cases, not for Random Forest. For the same algorithmic approach and neighbor (non) awareness, the models created with re-sampling scored higher in terms of AUCROC. The only exception was the (Random

Forest). This trend is consistent with the expectation and recent research on dealing with imbalanced datasets [163].

**Table 21. Average Results of Predicting Interpersonal Trust Using Various Classification and Sampling Methods.**

| Sampling Approach | Algorithmic Approach | Individual Path (Non-Neighbor-Aware) | | Neighbor-Aware | |
|---|---|---|---|---|---|
| | | Acc | AUCROC | Acc | AUCROC |
| **AS-IS** | **Random Forest** | 92.48% | 79.84% | 93.12% | 81.08% |
| **AS-IS** | **FC-DNN** | 98.63% | 47.53% | 95.04% | 71.07% |
| **AS-IS** | **NADAL** | 98.64% | 51.31% | 98.62% | 85.29% |
| **SMOTE+Tomek** | **Random Forest** | 93.78% | 77.85% | 94.67% | 78.94% |
| **SMOTE+Tomek** | **FC-DNN** | 98.63% | 49.51% | 92.59% | 86.35% |
| **SMOTE+Tomek** | **NADAL** | 92.54% | 90.63% | 94.55% | **93.23%** |

Next, we note that deep learning approaches (both FC-DNN and NADAL) generally outperform the shallow learning approach (Random Forest). Again, while we notice some exceptions, at least one of the two deep learning approaches always outperforms the shallow learning approach in each of the considered sampling and neighbor-awareness settings. This is again along expected lines as deep learning approaches tend to have more opportunity to capture linear and non-linear associations between different features and create comprehensive models.

Next, the neighbor-aware approach yields better performance in both shallow and deep machine learning approaches. Interestingly, while most algorithms struggled to learn the minority class well (AUCROC around 0.50) without SMOTE+Tomek resampling, NADAL approach was able to obtain credible performance with the neighbor-aware approach in even that setting. All comparisons between the same algorithmic approach

(Individual Path vs. Neighbor-Aware) were found to be statistically significant using two-tailed unpaired t-tests (at alpha= 0.05 level) suggesting that the *AUCROC of the Neighbor-Aware approaches were statistically significantly better than the Individual Path approaches.*

Lastly, we note that the NADAL approach worked better than the baseline deep learning approach (FC-DNN). This trend was especially obvious with the SMOTE+Tomek resampling i.e. in the setting where the performance was higher in general. This suggests that early fusion of features might not allow for the interrelationships within the same channel to be learned adequately without the influence of other channels. The step-wise unification of different channels across the architecture seems to have provided better opportunities for the social channels to learn both intra-channel and inter-channel relationships.

The highest overall AUCROC score of 93.23% was obtained using SMOTE+Tomek sampling, Neighbor-Aware features, and NADAL architecture. A score of 93.23% indicates that the model was able to learn both the majority and minority classes reasonably well and could be useful in practice where interpersonal trust needs to be inferred using phone based metadata.

## 4.6. Discussion

### 4.6.1. Methodological Considerations

The work presented here tackles the problem of inferring interpersonal trust automatically using phone metadata. Such a problem requires dealing with highly imbalanced datasets and also takes place in a socially rich setting. Hence, this work proposes and empirically tests the use of multiple techniques to improve the automatic prediction quality. While a SMOTE+Tomek approach allows better learning based on a balanced training set, the neighbor-aware approach allows for the use of neighboring connections' data for better inference. Finally, the growth in such data allows for the use of deep learning techniques to obtain better performance. However, the architectures for deep learning need to be defined in a manner that is responsive to the nature of the task at hand. In particular, the NADAL architecture, which allows for learning appropriate features from neighboring edges while also giving due credit to the primary node in question, was found to yield the best results. As the first effort in this direction, we have chosen to use Deep Neural Networks (DNN) using Artificial Neural Networks (ANN), which are relatively simple and well-studied in the deep learning literature. The positive results obtained here on utilizing neighboring relationships motivate the exploration of other techniques for future work.

**4.6.2. Privacy of User Data and Ethical Considerations**

The data for this study come from the MIT Friends and Family study [71], which has been adopted by multiple research groups to study questions pertaining to social and ubiquitous computing. We use a version of the dataset where all data was anonymized and hashed. Under no circumstance, the content of the Calls or SMS messages were available to the authors responsible for analysing the metadata. We recognize the ethical concerns related to automatic generation of scores to quantify human ties and their interpersonal trust. Similar concerns have been raised in the past for automatic generation of mental health scores for individuals or even survey based quantification of interpersonal relationships [121]. While waiting for the development of better privacy and ethics polices, we firmly believe that it takes various studies like this one to facilitate the comprehension around the visions of using ubiquitous data and enrich the discussion in the research community around them [139].

**4.6.3. Limitations**

The current study has some limitations. First is the relatively small sample size - 130 individuals. Hence, we are careful not to generalize the results until they're re-verified with a larger population of samples. Second is the homogeneity of the sample. While this limitation prevents us from generalizing the findings to larger populations, the homogeneity also allowed us to isolate socio-mobile behavior as a predictor. Another limitation is the use of a specific question-based trust metric in this work.

Despite these limitations, this study's value lies in the new ground it breaks in multiple ways. To our knowledge, there have been no previous studies undertaken that utilize a *deep learning approach* to infer interpersonal trust from sensor data. Correspondingly, the proposed NADAL architecture is the first deep learning based attempt at utilizing *neighbors' edge properties* to better infer aspects associated with a user's edge or relationships. We hope that the results obtained in this work will motivate more significant work which applies the abovementioned techniques to settings with diverse trust measurement methods and sampled populations.

**4.6.4. Implications**

With further validation, this line of research could have multiple implications for individuals as well as the society. The participants who opt-in to such automated interpersonal trust scoring apps could get better customized recommendations for social activities, news, and mobile commerce apps; whom one trusts is a critical mediator for almost all goods and services that one procures or exchanges online and offline. For instance, social media sites could recommend products that are rated high by the trusted peers. Similarly, trusted peers could be suggested for exchange of services. With enhancements, the proposed approach can be used to understand societal changes and support emerging socio-technical contexts like the sharing economy [123].

At a societal level, such apps could alleviate the need to run costly annual surveys to assess the trust-based “state of the nation” as proposed by [26]. Instead, automated methods can be used to create a real-time nation-wide trust census and make it a part of the public policy and decision making process. Further, an ability to study the phenomenon of interpersonal trust and its “in the wild” dynamics at scale can substantially advance the literature in several fields (e.g., economics, psychology) that study trust.

# Chapter 5

# Conclusion and Future Work

## 5.1. Conclusion

In this dissertation, we have proposed "personal big data" based new approaches to infer individual trust, altruism, and interpersonal trust using mobile phone features as an alternative to conventional methods like surveys and lab experiments. Using phone-based behavioral features allowed us to build predictive models by means of machine learning classification algorithms whose accuracy, AUC, and F1 scores were promising and encouraging. To the best of our knowledge, there have been no previous studies that analyze the link between individual trust and altruism propensity levels along with phone-based behavioral features while maintaining fairness in the classification process. Hence, these results pave way for more research on leveraging ubiquitous sensing data for understanding the interconnections between socio-mobile behavioral data and human behavioral propensities.

Also, we have proposed a new approach to automatically infer interpersonal trust via phone-based features using shallow and deep learning methods. This is the first effort to suggest and validate the use of behavioral features from *neighboring* relationships to better predict the interpersonal trust ties of the target relationship. The best results for this problem were obtained based on a novel deep learning architecture (NADAL), which efficiently uses neighboring relationship data.

With further technical and ethical ground work, the proposed approaches can be used for inferring human behaviors and propensities of individuals at a scale of billions of people. Hence, with the growth in mobile phone penetration, the proposed approaches could have multiple implications for individuals (e.g. customized applications) as well as societies as they engage in higher levels of technology-mediated interactions.

### 5.2. Future Work

The results obtained in this dissertation pave way for more research on leveraging ubiquitous sensing data for understanding the interconnections between socio-mobile behavioral data alongside trust propensity, altruism, and interpersonal trust using different methods and in varied settings. The proposed approaches can be enhanced in future work by including a larger number of participants, more detailed phone-based features, and considering larger time durations. While we undertake exploratory analysis to reduce the disparity in the prediction algorithms in terms of their performance for different genders in predicting trust, future work can undertake more detailed analysis based on different definitions of groups and define more sophisticated methods to increase fairness. There are also opportunities for improving the work by creating more advanced deep learning architectures that are also neighbor-aware.

Furthermore, the suggested phone-based methods could be expanded to study and predict other personal behaviors and traits such as gratitude, compassion, and happiness. Taken together such methods open ways to better model human beings based on ubiquitous sensing and act as a building block towards a healthier and happier society.

### 5.3. Dissertation Related Publications

This dissertation is based on the following publications:

[1] **G. F. Bati** and V. K. Singh, "Are You Altruistic? Your Mobile Phone Could Tell," in *The 3rd IEEE (IoP 2017)*, San Francisco, CA, 2017 [164].

[2] **G. F. Bati** and V. K. Singh, "“Trust Us”: Mobile Phone Use Patterns Can Predict Individual Trust Propensity," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. 330 [145].

[3] **G. F. Bati** and V. K. Singh, "Altrumetrics: Inferring Altruism Propensity Based on Mobile Phone Use Patterns," in *IEEE Trans. Big Data, 2018* [165].

[4] **G. F. Bati** and V. K. Singh, "A Neighbor-Aware Deep Learning Approach for Inferring Interpersonal Trust," submitted for publication.

[5] **G. F. Bati**, R. Inamdar, and V. K. Singh, " FAIRSTART: Towards Fairness in Reality Mining," submitted for publication.

## References

[1] C. Gurrin, A. F. Smeaton, and A. R. Doherty, “Lifelogging: Personal big data,” in *Foundations and Trends in Information Retrieval*, vol. 8, 2014, pp. 1–125.

[2] A. Ali, J. Qadir, R. Rasool, A. Sathiaseelan, A. Zwitter, and J. Crowcroft, “Big data for development: applications and techniques,” *Big Data Anal.*, vol. 1, no. 2, pp. 1–24, 2016.

[3] V. D. Blondel, A. Decuyper, and G. Krings, “A survey of results on mobile phone datasets analysis,” *EPJ Data Sci.*, vol. 4, no. 10, 2015.

[4] K. G. Tarakji *et al.*, “Using a novel wireless system for monitoring patients after the atrial fibrillation ablation procedure: the iTransmit study,” *Heart Rhythm*, vol. 12, no. 3, pp. 554–559, 2015.

[5] A. Decuyper, “On the research for big data uses for public good purposes Opportunities and challenges,” *Netcom*, vol. 30, no. 3–4, pp. 305–314, 2016.

[6] H. Grassegger and M. Krogerus, “The Data That Turned the World Upside Down,” *Vice Media*, vol. 18, Jan. 2017.

[7] G. Miller, “The smartphone psychology manifesto,” *Perspect. Psychol. Sci.*, vol. 7, no. 3, pp. 221–237, 2012.

[8] “What Is Altruism?,” *The Greater Good Science Center at UC Berkeley.* [Online]. Available: http://greatergood.berkeley.edu/topic/altruism/definition. [Accessed: 04-Apr-2017].

[9] J. A. Colquitt, B. A. Scott, and J. A. LePine, “Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships With Risk Taking and Job Performance,” *J. Appl. Psychol.*, pp. 909–927, 2007.

[10] R. Borum, “The Science of Interpersonal Trust,” *Ment. Health Law Policy Fac. Publ.*, Jan. 2010.

[11] C. M. Cheung and M. K. Lee, “Understanding consumer trust in Internet shopping: A multidisciplinary approach,” *J. Am. Soc. Inf. Sci. Technol.*, pp. 479–492, 2006.

[12] D. L. McGuinness, H. Zeng, P. P. D. Silva, L. Ding, D. Narayanan, and M. Bhaowal, “Investigations into Trust for Collaborative Information Repositories: A Wikipedia Case Study,” presented at the The World Wide Web Conference (WWW2006), Edinburgh, UK, 2006.

[13] M. J. Metzger, “Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research,” *J. Am. Soc. Inf. Sci. Technol.*, pp. 2078–2091, 2007.

[14] E. Sillence and P. Briggs, “Trust and engagement in online health a timeline approach,” *Handb PsycholCommun Technol*, vol. 33, pp. 469–87, 2015.

[15] K. Kelton, K. R. Fleischmann, and W. A. Wallace, “Trust in digital information,” *J. Am. Soc. Inf. Sci. Technol.*, pp. 363–374, 2008.

[16] M. Harding, B. Knowles, N. Davies, and M. Rouncefield, “HCI, Civic Engagement & Trust,” in *the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI ’15)*, New York, NY, USA, 2015.

[17] S. Okasha, “Biological Altruism (Stanford Encyclopedia of Philosophy),” Jul-2013. [Online]. Available: https://plato.stanford.edu/archives/fall2013/entries/altruism-biological/.

[18] J. Miranda *et al.*, “From the Internet of Things to the Internet of People,” *IEEE Internet Comput.*, vol. 19, no. 2, pp. 40–47, 2015.

[19] P. Barclay, “Trustworthiness and competitive altruism can also solve the ‘tragedy of the commons,’” *Evol. Hum. Behav.*, vol. 25, no. 4, pp. 209–220, 2004.

[20] F. Exadaktylos, A. M. Espín, and P. Brañas-Garza, “Experimental subjects are not different,” *Sci. Rep.*, vol. 3, 2013.

[21] T. Yamagishi and M. Yamagishi, “Trust and commitment in the United States and Japan. ‘Motivation and Emotion,’” vol. 18, no. 2, pp. 129–166, 1994.

[22] J. Giles, “Computational social science: Making the links,” *Nature*, vol. 488, no. 7412, pp. 448–450, 2012.

[23] V. K. Singh and R. R. Agarwal, “Cooperative phoneotypes: exploring phone-based behavioral markers of cooperation,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 646–657.

[24] C. Batson, *The altruism question: Toward a social-psychological answer*. New York: Psychology Press, 2014.

[25] A. Ben-Ner and F. Halldorsson, “Trusting and trustworthiness: What are they, how to measure them, and what affects them,” *J. Econ. Psychol.*, pp. 64–79, 2010.

[26] J. Ermisch, D. Gambetta, H. Laurie, T. Siedler, and S. C. Noah Uhrig, “Measuring people’s trust,” *J. R. Stat. Soc. Ser. A Stat. Soc.*, vol. 172, no. 4, pp. 749–769, 2009.

[27] A. Roy, C. Sarkar, J. Srivastava, and J. Huh, “Trustingness & trustworthiness: A pair of complementary trust measures in a social network,” in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, 2016.

[28] D. Gambetta, *Trust: Making and Breaking Cooperative Relations*. Cambridge, MA: Oxford and Basil Blackwell, 1988.

[29] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, “Not so different after all: A cross-discipline view of trust,” *Acad. Manage. Rev.*, pp. 393–404, 1998.

[30] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, “Sensing, understanding, and shaping social behavior,” *IEEE Trans. Comput. Soc. Syst.*, vol. 1, no. 1, pp. 22–34, 2014.

[31] D. H. McKnight and N. L. Chervany, “Trust and distrust definitions: One bite at a time,” *Trust Cyber-Soc.*, pp. 27–54, 2001.

[32] R. I. M. Dunbar, “Breaking Bread: the Functions of Social Eating,” *Adapt. Hum. Behav. Physiol.*, pp. 1–14, 2017.

[33] G. K. Sahi, H. S. Sekhon, and T. K. Quareshi, “Role of trusting beliefs in predicting purchase intentions,” *Int. J. Retail Distrib. Manag.*, pp. 860–880, 2016.

[34] C. H. S. Dwyer and K. Passerini, “Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace,” in *13th Americas Conference on Information Systems (AMCIS)*, Keystone, CO, 2007.

[35] S. West, A. Griffin, and A. Gardner, “Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection,” *J. Evol. Biol.*, vol. 20, no. 2, pp. 415–432, 2007.

[36] Y. Chu and H. Zhang, “Considering altruism in peer-to-peer internet streaming broadcast. 14th international workshop on Network and operating systems support for digital audio and video,” presented at the NOSSDAV ’04, Cork, Ireland, 2004.

[37] K. Farrahi and K. Zia, “Trust reality-mining: evidencing the role of friendship for trust diffusion,” *Hum.-Centric Comput. Inf. Sci.*, pp. 1–16, 2017.

[38] K. Zolfaghar and A. Aghaieb, “Evolution of trust networks in social web applications using supervised learning,” *World Congr. Inf. Technol.*, 2010.

[39] M. J. Kim, N. Chung, C.-K. Lee, and M. W. Preis, “Why do smartphone shoppers help others on websites? The effects of attachments on reciprocal altruism,” *Inf. Dev.*, vol. 32, no. 4, pp. 920–936, 2016.

[40] P. Hui, K. Xu, V. Li, J. Crowcroft, V. Latora, and P. Lio, “Selfishness, altruism and message spreading in mobile social network,” presented at the INFOCOM Workshops, Rio de Janeiro, Brazil, 2009.

[41] S. Hameed, A. Wolf, K. Zhu, and X. Fu, “Evaluation of Human Altruism Using a DTN-based Mobile Social Network Application,” presented at the Workshop Digitale Soziale Netze, Braunschweig, Germany, 2012, pp. 81–95.

[42] J. Berg, J. Dickhaut, and K. McCabe, “Trust, reciprocity, and social history,” *Games Econ. Behav.*, vol. 10, no. 1, pp. 122–142, 1995.

[43] J. Fogel and E. Nehmad, “Internet social network communities: Risk taking, trust, and privacy concerns,” vol. 25, no. 1, pp. 153–160, 2009.

[44] E. Magalhães, P. Costa, and M. J. Costa, “Empathy of medical students and personality: Evidence from the Five-Factor Model,” *Med. Teach.*, vol. 34, no. 10, pp. 807–812, Oct. 2012.

[45] R. D. Putnam, “Bowling alone: America’s declining social capital,” *J. Democr.*, vol. 6, no. 1, pp. 65–78, 1995.

[46] R. D. Putnam, “Tuning in, tuning out: The strange disappearance of social capital in America,” *Polit. Sci. Polit.*, vol. 28, no. 4, pp. 664–683, 1995.

[47] B. Lancee, “The economic returns of immigrants’ bonding and bridging social capital: The case of the Netherlands,” *Int. Migr. Rev.*, vol. 44, no. 1, pp. 202–226, 2010.

[48] H. Coffé and B. Geys, “Toward an empirical characterization of bridging and bonding social capital,” *Nonprofit Volunt. Sect. Q.*, vol. 36, no. 1, pp. 121–139, 2007.

[49] K. Newton, “Trust, social capital, civil society, and democracy,” *Int. Polit. Sci. Rev.*, vol. 22, no. 2, pp. 201–214, 2001.
[50] S. Valenzuela, N. Park, and K. F. Kee, “Is there social capital in a social network site?: Facebook use and college students’ life satisfaction, trust, and participation,” *J. Comput. Commun.*, vol. 14, no. 4, pp. 875–901, 2009.
[51] R. Putnam, “Social capital: Measurement and consequences,” *Can. J. Policy Res.*, vol. 2, no. 1, pp. 41–51, 2001.
[52] S. Yang, S. Kurnia, H. Lee, and S. Kim, “Mobile phones and social capital: A preliminary study in south Korea,” *PACIS 2008 Proc.*, p. 138, 2008.
[53] Y.-F. Chen, “Social Phenomena of Mobile Phone Use: An Exploratory Study in Taiwanese College Student,” in *the Second International Conference on Digital Communication Conference of the National Chung-Cheng*, University, Chia-Yi, Taiwan, 2006.
[54] V. K. Singh and I. Ghosh, “Inferring Individual Social Capital Automatically via Phone Logs,” *Press Proc. ACM Hum. Comput. Interact.*, vol. 1, no. 2, 2017.
[55] W. Z. Khan, Y. Xiang, M. Y. Aalsalem, and K. Arshad, “Mobile Phone Sensing Systems: A Survey,” *IEEE Commun. Surv. Tutor.*, vol. 15, no. 1, pp. 402–427, 2013.
[56] Y. A. de Montjoye, J. Quoidbach, F. Robic, and A. Pentland, “Predicting personality using novel mobile phone-based metrics,” in *6th International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, Washington, DC, 2013.
[57] Y. Guan and T. Plötz, “Ensembles of Deep LSTM Learners for Activity Recognition using Wearables,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 1–28, Jun. 2017.
[58] K. Krishna, D. Jain, S. V. Mehta, and S. Choudhary, “An LSTM Based System for Prediction of Human Activities with Durations,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–31, Jan. 2018.
[59] L. Peng, L. Chen, Z. Ye, and Y. Zhang, “AROMA: A Deep Multi-Task Learning Based Simple and Complex Human Activity Recognition Method Using Wearable Sensors,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 2, pp. 1–16, Jul. 2018.
[60] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, “Deep Learning for Sensor-based Activity Recognition: A Survey,” *ArXiv Prepr. ArXiv170703502*, 2017.
[61] N. Bantilan, “Themis-ml: A Fairness-Aware Machine Learning Interface for End-To-End Discrimination Discovery and Mitigation,” *J. Technol. Hum. Serv.*, vol. 36, no. 1, pp. 15–30, Jan. 2018.
[62] C. O’Neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, 2016.
[63] J. Buolamwini and T. Gebru, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification,” *Proc. 1st Conf. Fairness Account. Transpar. PMLR*, no. 81, pp. 77–91, 2018.
[64] B. Lepri, N. Oliver, E. Letouzé, A. Pentland, and P. Vinck, “Fair, Transparent, and Accountable Algorithmic Decision-making Processes,” *Philos. Technol.*, Aug. 2017.
[65] M. B. Zafar, I. Valera, M. Gomez Rodriguez, and K. P. Gummadi, “Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment,” in *Proceedings of the 26th International Conference on World Wide Web*, 2017, pp. 1171–1180.
[66] N. Eagle and A. S. Pentland, “Reality mining: sensing complex social systems,” *Pers. Ubiquitous Comput.*, vol. 10, no. 4, pp. 255–268, 2006.
[67] K. Cook, Ed., “Trust in society,” in *Trust in society*, New York, NY: Russell Sage Foundation, 2001.
[68] R. Wang *et al.*, “StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones,” in *the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, Washington, US, 2014.
[69] Y. A. de Montjoye, S. S. Wang, A. Pentland, D. T. Anh, and A. Datta, “On the Trusted Use of Large-Scale Personal Data,” *IEEE Data Eng Bull*, pp. 5–8, 2012.
[70] N. Lane, E. Miluzzo, H. Lu, D. Peebles, A. Campbel, and T. Choudhury, “A survey of mobile phone sensing,” *IEEE Commun. Mag.*, vol. 48, no. 9, pp. 140–150, Sep. 2010.
[71] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland, “Social fMRI: Investigating and shaping social mechanisms in the real world,” *Pervasive Mob. Comput.*, vol. 7, no. 6, pp. 643–659, 2011.

[72] A. Montoro, P.-C. Shih, M. Román, and A. Martínez-Molina, "Spanish adaptation of Yamagishi General Trust Scale," *An. Psicol. Psychol.*, vol. 30, no. 1, pp. 302–307, 2013.
[73] A. M. Evans and W. Revelle, "Survey and behavioral measurements of interpersonal trust," *J. Res. Personal.*, pp. 1585–1593, 2008.
[74] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo, "Socio-spatial properties of online location-based social networks," in *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)*, Barcelona, Spain, 2011, pp. 329–336.
[75] G. M. Vazquez-Prokopec *et al.*, "Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment," *PloS One*, vol. 8, no. 4, 2013.
[76] L. Canzian and M. Musolesi, "Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," presented at the The 2015 ACM international joint conference on pervasive and ubiquitous computing, Osaka, Japan, 2015.
[77] V. Kaufmann, M. M. Bergman, and D. Joye, "Motility: mobility as capital," *Int. J. Urban Reg. Res.*, vol. 28, no. 4, pp. 745–756, 2004.
[78] R. Oldenburg, *Celebrating the third place: Inspiring stories about the "great good places" at the heart of our communities*. Cambridge, MA: Da Capo Press, 2001.
[79] V. Kaufmann and E. D. Widmer, "Motility and family dynamics: Current issues and research agendas," *J. Fam. Res.*, vol. 18, no. 1, pp. 111–129, 2006.
[80] R. Oldenburg and D. Brissett, "The third place," *Qual. Sociol.*, vol. 5, no. 4, pp. 265–284, 1982.
[81] P. K. Jonasona, A. Jonesb, and M. Lyonsb, "Creatures of the night: Chronotypes and the Dark Triad traits," *Personal. Individ. Differ.*, vol. 55, no. 5, pp. 538–541, 2013.
[82] M. Lyons and S. Hughes, "Feeling me, feeling you? Links between the Dark Triad and internal body awareness," *Personal. Individ. Differ.*, pp. 308–311, 2015.
[83] A. Adan and H. Almirall, "Horne & Östberg morningness-eveningness questionnaire: A reduced scale," *Personal. Individ. Differ.*, vol. 12, no. 3, pp. 241–253, 1991.
[84] S. Saeb *et al.*, "Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study," *J. Med. Internet Res.*, vol. 17, no. 7, 2015.
[85] S. Abdullah, M. Matthews, E. L. Murnan, G. Gay, and T. Choudhury, "Towards circadian computing: early to bed and early to rise makes some of us unhealthy and sleep deprived," in *the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Seattle, Washington, 2014.
[86] W. H. Macey and B. Schneider, "The meaning of employee engagement," *Ind. Organ. Psychol.*, vol. 1, no. 1, pp. 3–30, 2008.
[87] N. S. Shami, M. Muller, A. Pal, M. Masli, and B. Geyer, "Inferring employee engagement from social media," in *the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015.
[88] J. Golbeck, Ed., *Computing with social trust*. New York: Springer Science & Business Media, 2008.
[89] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi, "Human Mobility, Social Ties, and Link Prediction," *17th ACM SIGKDD Int. Conf. Knowl. Discov. Data*, 2011.
[90] V. K. Singh, L. Freeman, B. Lepri, and A. P. Pentland, "Classifying spending behavior using socio-mobile data," vol. 2, no. 2, pp. 99–111, 2013.
[91] N. Lin, *Social capital: A theory of social structure and action*. New York, NY: Cambridge university press, 2002.
[92] R. Bapna, A. Gupta, R. Rice, and A. Sundararajan, "Trust and the strength of ties in online social networks: an exploratory field experiment," *MIS Q.*, vol. 41, no. 1, pp. 115–130, 2017.
[93] D. Williams, "On and Off the 'Net: Scales for Social Capital in an Online Era," *J. Comput.-Mediat. Commun.*, vol. 11, no. 2, pp. 593–628, 2006.
[94] M. S. Granovetter, "The Strength of Weak Ties," *Am. J. Sociol.*, vol. 78, no. 6, pp. 1360–1380, 1973.
[95] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," presented at the SIGCHI conference on human factors in computing systems, Boston, MA, 2009.
[96] R. E. Nelson, "The strength of strong ties: Social networks and intergroup conflict in organizations," *Acad. Manage. J.*, vol. 32, no. 2, pp. 377–401, 1989.
[97] D. Schulz, S. Bothe, and C. Körner, "Human mobility from gsm data - a valid alternative to gps," presented at the Mobile data challenge, 2012.

[98] V. K. Singh, B. Bozkaya, and A. Pentland, “Money walks: implicit mobility behavior and financial well-being,” *PloS One*, vol. 10, no. 8, p. e0136628, 2015.
[99] J. S. Coleman, “Social capital in the creation of human capital,” *Am. J. Sociol.*, vol. 94, pp. 95–120, 1988.
[100] L. Pappalardo, F. Simini, S. Rinzivillo, D. Pedreschi, F. Giannotti, and A. L. Barabási, “Returners and explorers dichotomy in human mobility,” *Nat. Commun.*, vol. 6, 2015.
[101] V. K. Singh, L. Freeman, B. Lepri, and A. S. Pentland, “Predicting spending behavior using socio-mobile features,” in *2013 International Conference on Social Computing*, 2013, pp. 174–179.
[102] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *J. R. Stat. Soc. Ser. B Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.
[103] R. C. Team, “R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014,” 2018.
[104] T. Hastie and B. Efron, “lars: Least Angle Regression, Lasso and Forward Stagewise,” 2013. [Online]. Available: https://CRAN.R-project.org/package=lars.
[105] M. Hall, “Correlation-based feature selection for machine learning,” Jan-1999. [Online]. Available: https://www.lri.fr/.
[106] E. Frank, M. A. Hall, and I. H. Witten, “The WEKA Workbench,” Jan-2016. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf.
[107] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: an update,” *ACM SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, 2009.
[108] O. D. Mining, “Orange Visual Programming Documentation.” Oct-2017.
[109] A. Zheng, *Evaluating Machine Learning Models A Beginner’s Guide to Key Concepts and Pitfalls*. Sebastopol, CA: O’Reilly Media, Inc, 2015.
[110] J. Neyman, “On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection,” *J. R. Stat. Soc.*, vol. 97, no. 4, pp. 558–625, 1934.
[111] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[112] A. Agarwal, A. Beygelzimer, M. Dudik, J. Langford, and H. Wallach, “A Reductions Approach to Fair Classification,” in *International Conference on Machine Learning*, 2018, pp. 60–69.
[113] M. Hardt, E. Price, and N. Srebro, “Equality of opportunity in supervised learning,” in *Advances in neural information processing systems*, 2016, pp. 3315–3323.
[114] F. PEDREGOSA, “Scikit-learn: Machine learning in python,” *J Mach. Learn Res*, vol. 12, pp. 2825–2830, 2011.
[115] D. Krackhardt, N. Nohria, and B. Eccles, “The strength of strong ties,” presented at the Networks in the knowledge economy, New York, 2003.
[116] M. Tanis and T. Postmes, “A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour,” *Eur. J. Soc. Psychol.*, vol. 35, pp. 413–424, 2005.
[117] M. J. Platow, M. Foddy, T. Yamagishi, L. Lim, and A. Chow, “Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership,” *Eur. J. Soc. Psychol.*, vol. 42, pp. 30–35, 2012.
[118] M. Foddy and R. Dawes, “Group-based trust in social dilemmas,” presented at the New Issues and Paradigms in Research on Social Dilemmas, New York, USA, 2008, pp. 57–85.
[119] L. K. Knobloch and D. H. Solomon, “Information Seeking Beyond Initial Interaction,” *Hum. Commun. Res.*, vol. 2, no. 28, pp. 243–257, 2002.
[120] V. Iosifidis and E. Ntoutsi, “Dealing with Bias via Data Augmentation in Supervised Learning Scenarios,” in *Proceedings of the International Workshop on Bias in Information, Algorithms, and Systems co-located with 13th International Conference on Transforming Digital Worlds (iConference 2018)*, Sheffield, United Kingdom, 2018, vol. 2103, pp. 1–6.
[121] Z. Tufekci, “Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency,” *Colo. Technol. Law J.*, vol. 13, no. 2, pp. 203–218, 2015.
[122] M. F. Cheunga and W. To, “The influence of the propensity to trust on mobile users’ attitudes toward in-app advertisements: An extension of the theory of planned behavior,” *Comput. Hum. Behav.*, vol. 72, pp. 102–111, 2017.

[123] M. Möhlmann and A. Geissinger, "Trust in the Sharing Economy: Platform-Mediated Peer Trust," in *Cambridge Handbook on the Law of the Sharing*, Economy, Cambrridge, UK: Cambridge University Press, 2018.

[124] T. Speicher *et al.*, "Potential for Discrimination in Online Targeted Advertising," in *Conference on Fairness, Accountability and Transparency*, 2018, pp. 5–19.

[125] B. Friedman, "Value-sensitive design," *interactions*, vol. 3, no. 6, pp. 16–23, 1996.

[126] S. Consolvo *et al.*, "Activity sensing in the wild: a field trial of ubifit garden," in *the SIGCHI Conference on Human Factors in Computing Systems*, 2008.

[127] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," presented at the The eighteenth annual ACM-SIAM symposium on Discrete algorithms, New Orleans, Louisiana, 2007, pp. 1027–1035.

[128] G. Danilo, S. MacDonald, and T. Archer, "Two different approaches to the affective profiles model: median splits (variable-oriented) and cluster analysis (person-oriented)," *PeerJ*, vol. 3, no. 1380, 2015.

[129] J. DeCoster, M. Gallucci, and A.-M. R. Iselin, "Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives," *J. Exp. Psychopathol.*, vol. 2, no. 2, pp. 197–209, 2011.

[130] D. H. Kim, N. K. Seely, and J.-H. Jung, "Do you prefer, Pinterest or Instagram? The role of image-sharing SNSs and self-monitoring in enhancing ad effectiveness. Computers in," *Hum. Behav.*, vol. 70, pp. 535–543, 2017.

[131] J. Rushton, R. Chrisjohn, and G. Fekken, "The altruistic personality and the self-report altruism scale," *Personal. Individ. Differ.*, vol. 2, no. 4, pp. 293–302, 1981.

[132] S. King Jr and M. J. Holosko, "The development and initial validation of the empathy scale for social workers," *Res. Soc. Work Pract.*, vol. 22, no. 2, pp. 174–185, 2012.

[133] J. -F. Le Galliard, R. Ferrière, and U. Dieckmann, "Adaptive Evolution of Social Traits: Origin, Trajectories, and Correlations of Altruism and Mobility," *Am. Nat.*, vol. 165, no. 2, pp. 206–224, 2005.

[134] O. Curry, S. Roberts, and R. Dunbar, "Altruism in social networks: Evidence for a 'kinship premium,'" *Br. J. Psychol.*, vol. 104, no. 2, pp. 283–295, 2013.

[135] T. M. Kodinariya and P. R. Makwana, "Review on determining number of Cluster in K-Means Clustering," *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, vol. 1, no. 6, pp. 90–95, 2013.

[136] L. Mouselimis, "ClusterR: Gaussian Mixture Models, K-Means, Mini-Batch-Kmeans and K-Medoids Clustering," *Retrieved*. [Online]. Available: https://cran.r-project.org/web/packages/ClusterR/. [Accessed: 11-Aug-2017].

[137] J. Demšar *et al.*, "Orange: Data Mining Toolbox in Python," *J. Mach. Learn. Res.*, vol. 14, pp. 2349–2353, 2013.

[138] J. Darley and B. Latane, *The unresponsive bystander: Why doesn't he help*. New York: Appleton-Century-Crofts, 1970.

[139] A. Shifali, J. Yttri, and W. Nilsen, "Privacy and security in mobile health (mHealth) research," *Alcohol Res. Curr. Rev.*, vol. 36, no. 1, pp. 143–151, 2014.

[140] T. Qin, W. Shangguan, G. Song, and J. Tang, "Spatio-Temporal Routine Mining on Mobile Phone Data," *ACM Trans Knowl Discov Data*, vol. 12, no. 5, pp. 56:1–56:24, Jun. 2018.

[141] H. M. S. Hossain, M. A. Al Haiz Khan, and N. Roy, "DeActive: Scaling Activity Recognition with Active Deep Learning," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 2, pp. 1–23, Jul. 2018.

[142] F. Bonchi, "Influence Propagation in Social Networks: A Data Mining Perspective," in *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01*, 2011, p. 2.

[143] R. Bapna and A. Umyarov, "Do your online friends make you pay? A randomized field experiment on peer influence in online social networks," *Manag. Sci.*, vol. 61, no. 8, pp. 1902–1920, 2015.

[144] M. S. H. Mukta, M. E. Ali, and J. Mahmud, "Identifying and validating personality traits-based homophilies for an egocentric network," *Soc. Netw. Anal. Min.*, vol. 6, no. 1, p. 74, Dec. 2016.

[145] G. F. Bati and V. K. Singh, "'Trust Us': Mobile Phone Use Patterns Can Predict Individual Trust Propensity," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. 330.

[146] S. Adali, *Modeling trust context in networks*. Springer, 2013.

[147] F. Yakoub, M. Zein, K. Yasser, A. Adl, and A. E. Hassanien, "Predicting personality traits and social context based on mining the smartphones SMS data," in *Intelligent Data Analysis and Applications*, Springer, 2015, pp. 511–521.

[148] L. Jin, S. Gao, Z. Li, and J. Tang, "Hand-Crafted Features or Machine Learnt Features? Together They Improve RGB-D Object Recognition," in *2014 IEEE International Symposium on Multimedia*, Taichung, Taiwan, 2014, pp. 311–319.

[149] W. Li, S. Manivannan, S. Akbar, J. Zhang, E. Trucco, and S. J. McKenna, "Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, Czech Republic, 2016, pp. 1405–1408.

[150] T. Majtner, S. Yildirim-Yayilgan, and J. Y. Hardeberg, "Combining deep learning and hand-crafted features for skin lesion classification," in *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu, Finland, 2016, pp. 1–6.

[151] G. Antipov, S.-A. Berrani, N. Ruchaud, and J.-L. Dugelay, "Learned vs. Hand-Crafted Features for Pedestrian Gender Recognition," in *Proceedings of the 23rd ACM international conference on Multimedia - MM '15*, Brisbane, Australia, 2015, pp. 1263–1266.

[152] S. Greenspan, D. Goldberg, D. Weimer, and A. Basso, "Interpersonal Trust and Common Ground in Electronically Mediated Communication," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, New York, NY, USA, 2000, pp. 251–260.

[153] G. Lemaıtre and F. Nogueira, "Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning," p. 5.

[154] G. E. Batista, A. L. Bazzan, and M. C. Monard, "Balancing Training Data for Automated Annotation of Keywords: a Case Study.," in *WOB*, 2003, pp. 10–18.

[155] N. CHAWLA, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[156] "Techniques to deal with imbalanced data | Kaggle." [Online]. Available: https://www.kaggle.com/npramod/techniques-to-deal-with-imbalanced-data. [Accessed: 11-Aug-2018].

[157] G. Csardi and T. Nepusz, "The igraph software package for complex network research. InterJournal Complex Systems 1695," *Available Igraph Org Accessed Novemb.*, vol. 30, p. 2015, 2006.

[158] K. Ognyanova, "Static and dynamic network visualization with R." [Online]. Available: http://kateto.net/network-visualization. [Accessed: 11-Aug-2018].

[159] V. BUSKENS, W. RAUB, and J. VAN DER VEER, "Trust in triads: An experimental study," *Soc. Netw.*, vol. 32, no. 4, pp. 301–312, 2010.

[160] D. Dang-Pham, S. Pittayachawan, and V. Bruno, "Applying network analysis to investigate interpersonal influence of information security behaviours in the workplace," *Inf. Manage.*, vol. 54, no. 5, pp. 625–637, Jul. 2017.

[161] V. Radu *et al.*, "Multimodal deep learning for activity and context recognition," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 4, p. 157, 2018.

[162] F. Chollet and others, "Keras," 2015. [Online]. Available: https://keras.io.

[163] N. V. Chawla, "DATA MINING FOR IMBALANCED DATASETS: AN OVERVIEW," *DATA Min. Knowl. Discov. Handb.*, p. 15.

[164] G. F. Bati and V. K. Singh, "Are you altruistic? Your mobile phone could tell," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 2017.

[165] G. F. Bati and V. K. Singh, "Altrumetrics: Inferring Altruism Propensity Based on Mobile Phone Use Patterns," *IEEE Trans. Big Data*, pp. 1–1, 2018.