PERSONALIZED ESTIMATION OF DAILY EMOTIONS AND INTERPERSONAL CONFLICT BETWEEN ROMANTIC PARTNERS VIA METRIC LEARNING

A Thesis

by

VARSHA VENKATARAMU

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Dileep Kalathil
Committee Members,	Theodora Chaspari
	Stavros Kalafatis
Head of Department,	Miroslav M. Begovic

December 2021

Major Subject: Computer Engineering

ABSTRACT

A novel personalized model using metric learning via siamese Neural Network is implemented to estimate daily emotions and detect interpersonal couples' conflict using moment-to-moment multimodal bio-behavior signals (i.e., physiological, linguistic and acoustic signals) and additional relationship characteristics that includes individuals' relationship satisfaction and attachment information. The ambulatory couples' data has high inter-participant variability because each participant have a different distribution of data. Hence, a personalized model that has ability to eliminate the inter-participant variability and preserve the behavior characteristic is likely to perform well. Personalized learning implemented using metric learning via siamese neural network have innate ability to rank the pair of inputs after learning the personalized embeddings. Variants of the proposed personalized model and loss functions have been implemented and explored in this study. The performance of these proposed models are compared against that of non-personalized models such as feed-forward neural network.

DEDICATION

To my mother, father, grandfather, and grandmother.

ACKNOWLEDGMENTS

I am immensely grateful to my advisor, Dr. Theodora Chaspari for providing me an opportunity to conduct research under her guidance. I am extremely thankful for her patience throughout my thesis study. Her research knowledge, valuable feedback and continuous support greatly helped me to complete this study.

I would like to express my deepest appreciation to my committee members, Dr. Dileep Kalathil and Dr. Stavros Kalafatis for their insightful feedback, time and patience.

I would like to thank my fellow researcher Projna Parojamita at HUman Bio-Behaviour Signals (HUBBS) lab for her valuable assistance, feedback and continuous support during my thesis study.

I would like to thank the Texas A&M University for providing me an opportunity to pursue Master of Science Degree in Computer Engineering.

I would like to thank my parents for supporting and guiding me throughout. Finally, I would like to thank my brother and my dearest friends, whose constant encouragement helped me in the completion of the study.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

The thesis committee for this work includes Dr. Dileep Kalathil (Chair) of the Department of Electrical and Computer Engineering at Texas A&M University, Dr. Theodora Chaspari (Member) of the Department Computer Science at Texas A&M University and Dr. Stavros Kalafatis(Member) of the Department of Electrical and Computer Engineering.

The data utilized for the research was provided by Dr. Gayla Mangolin of Department of Psychology at University of Southern California.

All other work conducted as part of this thesis was completed by the author independently.

Funding Sources

Graduate study was performed without utilizing any form of funding sources.

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iii

ACKNOWLEDGMENTS iv

CONTRIBUTORS AND FUNDING SOURCES v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES ix

1. INTRODUCTION, LITERATURE REVIEW AND PRACTICAL APPLICATIONS 1

1.1 Introduction 1
1.2 Literature Review 1
1.3 Practical Application 2

2. RESEARCH QUESTIONS 4

3. METHODOLOGY 6

3.1 Data description and processing 6
3.1.1 Data collection methods 7
3.1.2 Types of features and processing 7
3.1.3 Self-assessments 8
3.1.4 Data post-processing 9
3.1.5 Data summary 10
3.2 Siamese neural network (SNN) 10
3.2.1 Fundamentals of SNN 11
3.2.2 Regression & classification 12
3.2.2.1 Regression 12
3.2.2.2 Classification 14
3.2.3 Forming pairs of samples for regression/classification 16
3.2.4 Hyper-parameter tuning & stratified cross-validation 18
3.2.4.1 Stratified cross-validation 18
3.2.4.2 Hyper-parameter tuning 20
3.3 Methods of personalization 20

3.3.1 SNN for achieving personalization ... 21
3.3.2 Personalization via learning from the same participant (as in the test set) ... 22
3.3.3 Individuals' relationship characteristic features as additional input to the SNN and in the last layer ... 23
3.4 Non-personalized Model ... 26

4. RESULTS ... 27

5. DISCUSSION ... 40

6. SUMMARY, LIMITATIONS AND FUTURE WORK ... 43
6.0.1 Summary ... 43
6.0.2 Limitations ... 43
6.0.3 Future Work ... 43

REFERENCES ... 45

LIST OF FIGURES

FIGURE Page

3.1 Outline of siamese neural network (SNN) architecture 12

3.2 Siamese neural network (SNN) architecture implementing regression. 13

3.3 Siamese neural network (SNN) architecture implementing classification. 15

3.4 Stratified cross-validation 19

3.5 Personalization via SNN 22

3.6 Introduce participant data to model prior to test phase 24

3.7 Individuals' relationship characteristic features as additional input to SNN 25

3.8 Individuals' relationship characteristic features added only to the last layer........... 25

3.9 Non-personalized Model 26

4.1 Experiments summary 29

4.2 Personalized Models without Relationship Characteristics: Regression 30

4.2 Personalized Models without Relationship Characteristics:Regression 31

4.3 Personalized Models with Relationship Characteristics on first layer: Regression 32

4.3 Personalized Models with Relationship Characteristics on first layer: Regression 33

4.4 Personalized Models with Relationship Characteristics on last-layer: Regression 34

4.4 Personalized Models with Relationship Characteristics on last-layer: Regression 35

4.5 Non-personalized Models with Relationship Characteristics: Regression 36

4.5 Non-personalized Models with Relationship Characteristics: Regression 37

4.6 Non-personalized Models without Relationship Characteristics: Regression.......... 38

4.6 Non-personalized Models without Relationship Characteristics: Regression.......... 39

5.1 Individuals' relationship characteristic features correlation with outcomes............ 42

LIST OF TABLES

TABLE Page

3.1 Ethnicity information of the participants 6

3.2 Career information of the participants 6

3.3 Input features summary post-processing 10

3.4 Hyper-parameters 20

1. INTRODUCTION, LITERATURE REVIEW AND PRACTICAL APPLICATIONS

1.1 Introduction

Interpersonal relationships are an integral part of every person's life and play a major role in one's physical and mental well-being [1]. There are mainly five types of interpersonal relationships, namely, friendship, romantic, platonic, family, and work relationships. Research suggests that an individual surrounded by satisfying relationships is likely to feel happy and protected, be productive, and have career success [2] [3]. In short, it is a human need to form interpersonal relationships, and therefore having strong and caring relationships increases one's chances to remain healthy and live longer. Amongst all forms of relationships, the depth of emotional connection and amount of quality time invested in romantic relationships is of considerate amount. Approximately 69% of the U.S. population are involved in romantic relationships or are looking forward to one [4], therefore rendering this topic a prevalent one to study. The goal of this M.S. thesis is to designing novel machine learning algorithms for quantifying emotions between romantic partners and detecting instances of interpersonal conflict using ambulatory signal-based data obtained from the partners in real-life. Successful completion of this study has the potential to provide a path for enhancing our understanding regarding micro-level behaviors between romantic partners in real life and facilitating the early diagnosis of challenges in interpersonal relationships.

1.2 Literature Review

Ambulatory data refer to data gathered in a real world environment by wearable and mobile sensors. These can be used to assist an individual in various aspects of their life and quantify facets of mental and emotional health [5]. Despite these opportunities, the analysis of ambulatory data presents significant challenges. These include the large amount of noise present in ambulatory environments, the unstructured nature of tasks performed by the participants being monitored, as well as the high inter-individual variability inherently present among participants [6].

Personalized learning consists of models that learn data patterns for each individual with the goal to accurately detect and predict the considered outcomes. Personalized models [7] and group-specific models [8] have been the focus of many recent studies due to their ability to yield reliable performance and provide improved results compared to general non-personalized models. Personalized models have the potential to eliminate the noise and inter-individual variability among participants in the collected data.

Metric learning approaches can naturally rank input pairs based on their similarity. This unique factor has helped personalized models implemented using metric learning via siamese neural network to provide good results in multiple classification tasks such as one-shot learning [9], visual tracking [10], speech based emotion recognition [11], gait-based user identification [12], and prediction of continuous progressive outcomes [13]. In addition to classification tasks, metric learning implemented via siamese neural network perform well in regression tasks in 3D object pose estimation [14]. Grounded in these findings, this thesis aims to explore the potential of using metric learning for designing personalized models of couples' well-being.

In this thesis work, a metric learning approach implemented via siamese neural networks (SNN) is designed to detect couples' emotions and interpersonal conflict using multimodal ambulatory data. The proposed metric learning approach is formulated as both a classification and a regression task. The results of proposed method of personalization are compared with those from non-personalized models that learn generic behavior patterns from the data and are implemented via a feed-forward neural network. The data used for the study are collected as part of USC Couple Mobile Sensing Project [15], including 87 couples and a total of 1560 samples.

1.3 Practical Application

Psychological intervention techniques are used widely in applied psychology to bring change in peoples behavior, promote good mental health and develop a habit [16]. For example, psychological intervention technique used on people to maintain relationship satisfaction and empathy in romantic relationship [17]. Positive psychology intervention technique focuses on increasing positive thoughts and emotions [18]. These positive psychology interventions are widely used in multiple fields such as relationship counseling, life coaching, mindfulness etc. Moment-to-moment feedback of psychological interventions at different decision stages are used to guide an individual towards emotional change [19].

Around 55.2% of adults in the US with mental health condition, have not received relevant mental health services in the year 2019 [20]. Mental health services are made easily accessible and affordable to individuals through smartphones and technology. Just-in-time adaptive psychological interventions are built using mobile technology, which provide right type and right amount of support to an individual at the right time [21]. The pragmatic framework of just-in-time adaptive interventions have been used as mobile health solutions for addictive behaviors [22] and to reduce sedentary behavior in obese adults [23].

The proposed personalized model in this thesis work has a probable application in a pragmatic framework of just-in-time adaptive psychological interventions amongst romantic partners participating in relationship counselling.

2. RESEARCH QUESTIONS

This M.S. thesis aims to answer the following research questions:

- **RQ1: To what extent can metric learning approaches learn personalized embeddings of multimodal data for the purpose of detecting emotion-based outcomes and interpersonal conflict?**

  This work quantifies various daily emotional states and also detect conflict amongst couples in a personalized manner by studying the bio-behavioral features (i.e., physiological, acoustic, linguistic) of both partners. This objective is carried out by building a personalized machine learning model, which uses metric learning approach implemented via Siamese neural network(SNN). The performance of SNN is compared to that of non-personalized models, such as feed-forward neural network, which model general non-personalized representations of the considered data. We hypothesize the metric learning approaches will outperform distribution-based learning, since they rely on modeling relative distances between the different levels of a given outcome, therefore are likely to require less data for training.

- **RQ2: To what extent integrating participants' individual relationship characteristics of participants yields improved performance?**

  The second objective of this work is to integrate partners' relationship characteristics (i.e., satisfaction, attachment) to the model during training. We hypothesize that integrating information about participants' relationship characteristics, in addition to the moment-to-moment multimodal information, will improve the overall model performance.

- **RQ3: To what extent integrating data samples from a target participant contributes to improved performance?**

  The third objective of this work is to utilize a portion of the target participants' data sample to train the machine learning models, so that the latter can learn personalized patterns explicitly for each individual. Our hypothesis is that by doing so, we will obtain improved results compared to not including data samples from a target participant.

3. METHODOLOGY

3.1 Data description and processing

The data used for the study is collected as part of the University of Southern California (USC) Couple Mobile Sensing Project [15]. The study is conducted with 87 couples aged between 18-25 years old, aimed to include young adults having unique and adverse childhood experience [8]. This contributes toward studying the impact of childhood background on the current romantic relationship, as well as its long-term implications on relationships. The data is diverse in terms of ethnicity, career, and stage of romantic relationship. Table 3.1 and 3.2 depicts the diversity in terms of ethnicity and career. On an average, the couples had been dating from 29.2 ± 24.2 months and 43.7% of couples were cohabiting.

Ethnicity	Percentage of participants(%)
Caucasian	27
Hispanic/Latino	25.9
African American	16.7
Asian	12.6
Multiracial	13.2
Other	4.6

Table 3.1: Ethnicity information of the participants

Career	Percentage of participants(%)
Working professionals	27
Part-time students	39.35
Full-time students	11.75

Table 3.2: Career information of the participants

3.1.1 Data collection methods

Each participant were provided with smartphones, wearable devices and sensors to collect data for one day, with data collected on regular intervals [8]. The data collection was performed from 9am till the end of the day. The Nexus 5 phone was used to collect GPS coordinates and 3-min audio samples every 12 minute. Physiological sensors namely Actiwave sensors[24] and Q sensors[25] was used in data collection. Actiwave sensor was used to collect electrocardiogram signal (ECG) whereas Q sensor was used to collect Electrodermal activity (EDA), wrist acceleration, body temperature, etc.

3.1.2 Types of features and processing

Moment-to-moment multimodal bio-behavioral features (i.e., physiological, acoustic, linguistic) of partners were used for this study [8]. An individuals current state of emotions directly affects moment-to-moment multimodal bi-behavioral features and hence these features are considered as primary inputs in the study.

Brief description of the moment-to-moment multimodal bio-behavioral features are as follows:

- **Physiological features** for both partners were extracted from electrodermal activity (EDA), electrocardiogram (ECG), body activity, body temperature. EDA features include mean skin conductance level and skin conductance responses(SCR). The SCR features are namely number, frequency and amplitude of SCR with threshold of 0.01 and 0.02 muS. ECG features are inter beat interval(IBI), heart rate(mean, min, max, standard deviation per minute), heart rate variability and R-R interval(mean, standard deviation). Body activity is the l^2 norm of 3-axis acceleration signal.

- **Linguistic features** were collected from the audio samples using Linguistic Inquiry and Word Count Dictionary (LIWD) [26]. The audio recordings from the audio samples were used to generate a manual transcript. LIWD extracts words from the transcript which describe the emotional and cognitive state of an individual. These words are categorized into linguistic (e.g., personal pronouns, verbs), psychological (e.g., positivity, negativity, swearing), personal concern (e.g., home-, work-, health related words), spoken categories (i.e para-linguistic features).

- **Acoustic features** include mean, median, maximum, minimum, standard deviation, and range for every hour of audio signals frequency and loudness.

3.1.3 Self-assessments

The couples participated in the study were provided with Ecological Momentary Assessments(EMA) every hour throughout the study. EMA's were provided on smartphones to each individual on hourly basis, for reporting mood, quality of interaction, and interpersonal conflict. Emotions for each partner such as anger, happiness, sadness, nervousness, stress, closeness as well as the conflict labels were self reported through EMA.

The couples prior to participation in the study were provided with surveys to complete, named the Quality of Marriage Index (QMI) [27] and Experiences in Close Relationships-Revised (ECR-R) [28].

- QMI contains questions related to various aspects of relationship, first five question on partners view of the relationship satisfaction in terms of various aspects of relationship and the last question on general relationship satisfaction. Relationship satisfaction are directly related to partners' happiness, stress and sadness etc. In psychology literature, there exists study of relationship satisfaction having an implication on conflict, congruence, empathy, unconditional regard and self-esteem [29] [30].

- ECR-R data is an 18-item questionnaire related to comfort and security in the relationship, therefore it captures the avoidance and anxious attachment data. Relationship anxiety refers to fear which arises due to questions on relationship stability, partners love and commitment. This fear is likely to raise anger arguments amongst partners or sad/unhappy thoughts in an individual. Avoidance refers to comfort level in the relationship to share private thoughts that provide feeling of closeness amongst partners. Sharing private thoughts is also likely to create a stress-free environment for partners and improve understanding amongst couples. Good understanding amongst partners has an impact on happiness of partners. In psychology literature, there exists a study of different relationship attachment styles having an implication on partners' behavior and emotion [31].

Therefore, data extracted from ECR-R and QMI assessments are also referred to as relationship characteristic features, and they are used as additional inputs features in the study.

Therefore to summarize the data extracted from the self-assessments:

- The **ground truth** emotions of both the partners are anger, happiness, sadness, nervousness, stress, closeness and interpersonal conflict extracted from self-reported assessments EMA.
- The individuals' **relationship characteristic features** of the partners are extracted from self-reported QMI and ECR-R survey. Relationship characteristic features include anxious and avoidance attachment extracted for both the partners from ECR-R. These features also include relationship satisfaction data from QMI, i.e., the mean value of first five questions of QMI and value of the last question of QMI for both the partners. These features are used as additional inputs along with the moment-to-moment multimodal bio-behavior features in some of the experiments in this study.

3.1.4 Data post-processing

Since the data was collected hourly the entire day for each participant (i.e. couples), there are multiple entries for each participant. Missing data in the moment-to-moment mutimodal bio-behavior features were filled with participant-wise mean whereas participants with missing relationship characteristic features were removed. Also, the participants that have missing self-reported ground truth (i.e angry, happy, sad, nervous, stress, close, conflict for both partners) were discarded. Feature-wise min-max normalization was performed on the entire input features.

Hence after the data cleaning, the total number of input features is listed in the table 3.3, the total number of participants (couples) are 39, the total number of self-reported ground truths are 13 out of which are 12 daily emotions for both partners(i.e. angry, happy, sad, nervous stress, close for both partners) and conflict. The relationship characteristics features are available only for 39 participants (couples). Hence post-processing the data results in a reduced number of participants i.e. 39 couples.

Input features	Total (includes both partners)
Physiological Features	24
Linguistic Features	110
Acoustic Features	54
Relationship Characteristic Features	8

Table 3.3: Input features summary post-processing

3.1.5 Data summary

Therefore, participants [15] moment-to-moment multimodal bio-behavior features (i.e., physiological, acoustic and linguistic) of both partners are used as primary inputs. Additionally, relationship characteristic features (i.e., QMI and ECR-R data) are used as additional inputs along with multimodal features.

The self-reported emotions such as anger, happiness, sadness, nervousness, closeness and stress by both the partners are considered to be ground-truth for the study. In addition to self-reported emotions, the self-reported conflict labels are also treated as ground-truth.

Predicting daily emotions are regression task whereas predicting interpersonal conflict is a classification task.

3.2 Siamese neural network (SNN)

The data has high inter-participant variability because each participant will have a different distribution. Hence, a model that has ability eliminate the inter-participant variability and preserve the behavior characteristic is likely to perform well. Therefore, metric learning based personalization is utilized for the study. Here, a SNN based model is used for building the above mentioned metric learning model.

3.2.1 Fundamentals of SNN

Siamese neural network were initially introduced for signature verification in an image matching problem during early 1990s [32]. It has two identical neural networks, which share the same weights. Siamese neural network needs two inputs, these two inputs can be identical or non-identical. SNN learns from the inputs such that it can differentiate between identical and non-identical inputs. The main idea behind sharing weights in the twin-network of SNN is to build a metric learning model with a loss function that minimizes the loss value upon identical inputs and maximizes the loss value upon non-identical inputs.

Figure 3.1 represents a basic SNN. The SNN has a twin-network structure where each of them have L hidden layers and all weights between these L layers are shared between the twin-networks. The distance between the last layers of the twin-network is used in the learning, which categorises this network as a metric learning procedure. Metric learning is a domain in supervised machine learning which focuses on learning from distance or similarity between inputs rather than the input objects itself. The symmetry and consistency property of the twin-network structure of SNN ensures that the pair of inputs are inherently ranked [33].

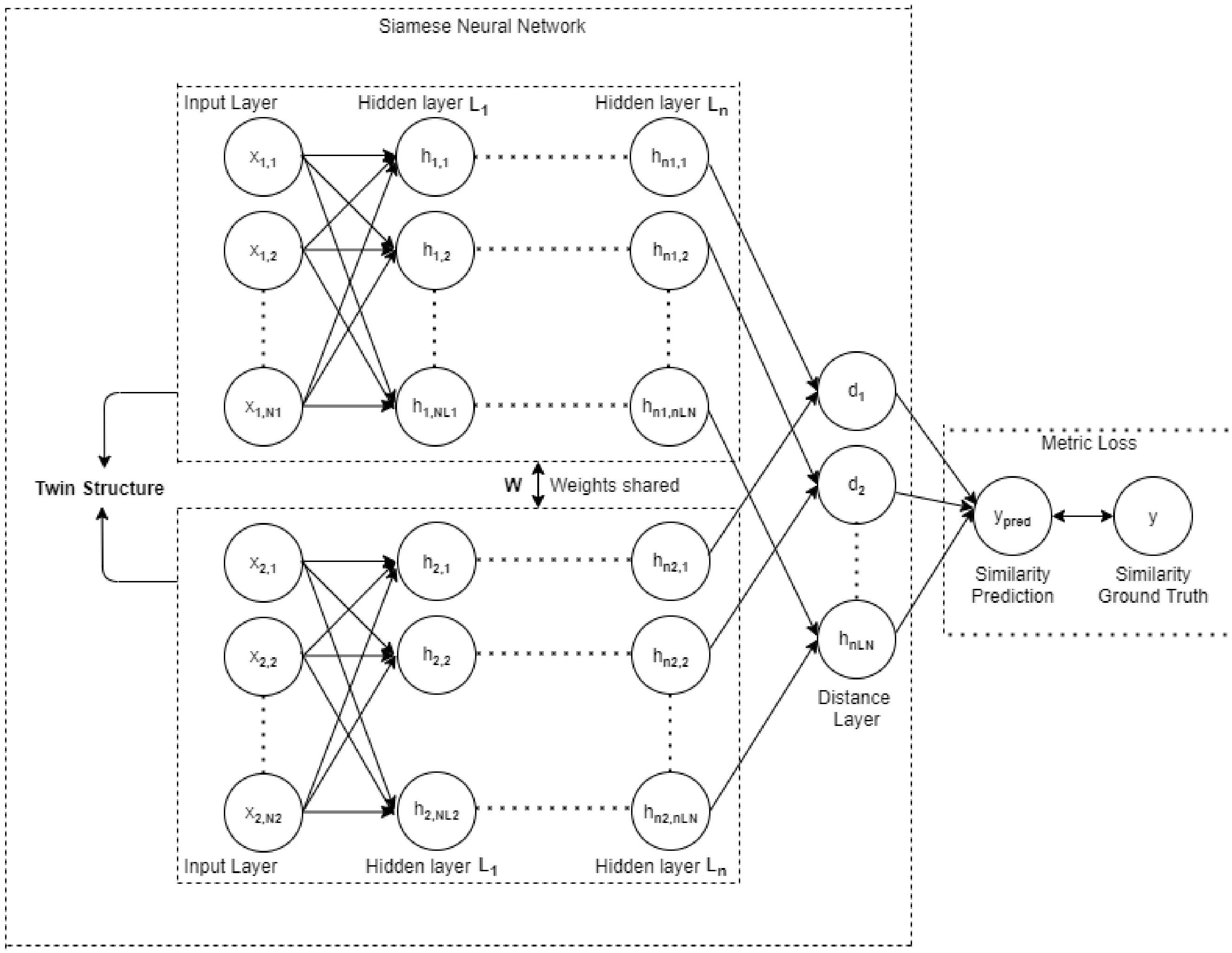


Figure 3.1: Outline of siamese neural network (SNN) architecture

3.2.2 Regression & classification

3.2.2.1 Regression

Ground truth daily emotions of both partners (i.e. angry, close, happy, nervous, sad and stress) is a regression task. The outcomes are in the range of 0-100. Siamese Neural Network is used as follows to build a personalized model for this regression task.

The figure 3.2 depicts the regression model for metric learning via SNN. Let x_1 and x_2 be the two inputs of the Siamese twin-structure. Let W be the transformation matrix of the twin-identical structure, since the weights are shared. The transformed inputs following the twin structure of Siamese are $W(x_1)$ and $W(x_2)$. The twin-structure outcomes are separately fed to two different feed-forward neural network for predicting the regression outcome of inputs x_1 and x_2. Let $F_1(W(x_1)$ and $F_2(W(x_2)$ be the outcomes of feed-forward neural network (FNN). The feed-forward neural network (FNN) in the fig 3.2 is one hidden layer with ReLu activation. Let y_1 and y_2 be one of the daily emotion considered as an outcome for training the model. The metric of twin-structure outcomes must be as close as possible to the difference between the outcomes of the inputs x_1 and x_2 i.e. $S = y_1 - y_2$. Let S' be the predicted outcome of the twin-structure and y_1' and y_2' be the predicted outcome of the two feed-forward regression networks.

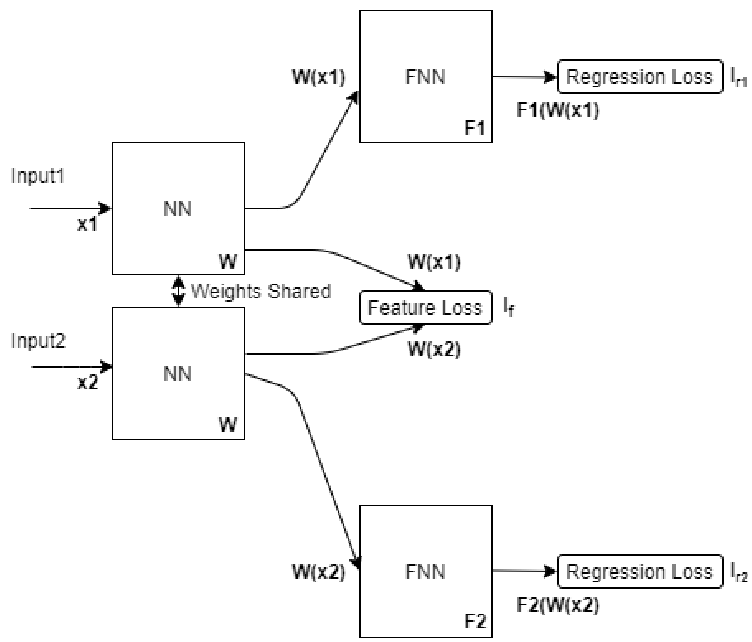


Figure 3.2: Siamese neural network (SNN) architecture implementing regression.

Therefore, metric loss for the above defined regression problem is named as feature loss l_f. Feature loss is defined as in Equation 3.1, where N is the total number of possible pair of inputs.

$$l_f = \sum_N \|W(x_1) - W(x_1)\|_2^2 - \|y_1 - y_2\|_2^2 \tag{3.1}$$

The regression loss of the two feed-forward networks are defined as l_{r_1} and l_{r_2} in Equation 3.2 and 3.3 using mean squared loss function, where N is the total number of possible pair of inputs.

$$l_{r_1} = \frac{1}{N} \sum_{N} \|F_1 - y_1\|_2^2 \tag{3.2}$$

$$l_{r_2} = \frac{1}{N} \sum_{N} \|F_2 - y_2\|_2^2 \tag{3.3}$$

Total loss function L of the designed siamese based network for the regression problem is sum of both feature loss l_f and regression loss l_{r_1} and l_{r_2}. Therefore $L = l_f + l_{r_1} + l_{r_2}$. The ground truths are unit normalized to avoid over-flow in the loss computation.

All the parameters of the regression model W, F_1 and F_2 are learnt in the training process such that the total loss L is as minimum as possible.

3.2.2.2 Classification

Ground truth conflict is a binary classification task. The value 1 represents conflict and value 0 represents non-conflict. Siamese Neural Network is used as follows to build a personalized model for this classification task.

The figure 3.3 depicts the classification model for metric learning via SNN. Let x_1 and x_2 be the two inputs of the Siamese twin-structure. Let W be the transformation matrix of the twin-identical structure, since the weights are shared. The transformed inputs following the twin structure of Siamese are $W(x_1)$ and $W(x_2)$. The twin-structure outcomes are separately fed to two different feed-forward neural network for predicting the classification outcome of inputs x_1 and x_2. Let $F_1(W(x_1)$ and $F_2(W(x_2)$ be the outcomes of feed-forward neural network (FNN). The feed-forward neural network (FNN) in the fig 3.3 is one hidden layer with sigmoid activation. Let y_1 and y_2 be conflict considered as an outcome for training the model. The metric of twin-structure outcomes must be as close as possible to the similarity between the outcomes of the inputs x_1 and x_2 defined as S in Equation 3.4. Let S' be the predicted outcome of the twin-structure and y_1' and y_2' be the predicted outcome of the two feed-forward classification networks.

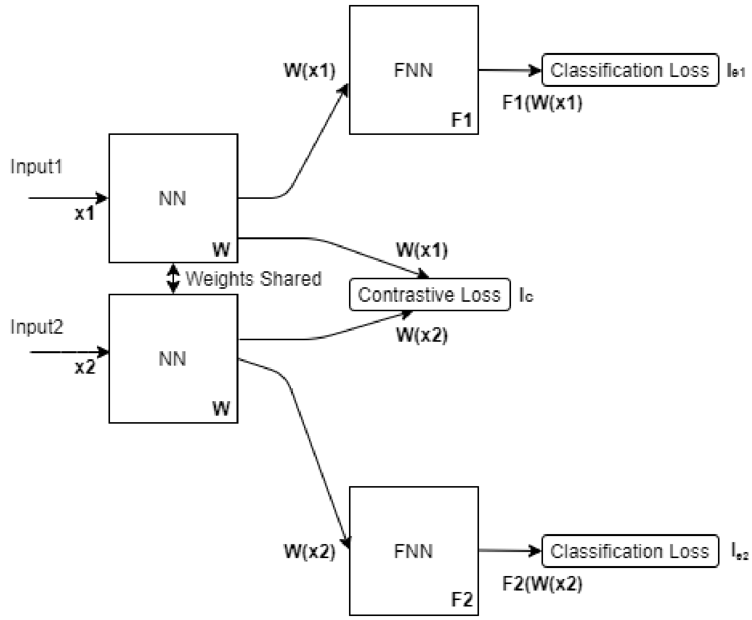


Figure 3.3: Siamese neural network (SNN) architecture implementing classification.

$$S = \begin{cases} 1 & \text{if } y_1 == y_2 \\ 0 & \text{if } y_1 != y_2 \end{cases} \tag{3.4}$$

Therefore, metric loss for the above defined classification problem is named as contrastive loss l_c. Contrastive loss is defined as in Equation 3.5, where N is the total number of possible pair of inputs and m is the margin. The contrastive loss is defined such that that the similar input pairs are pushed to have outcome distance close to 0 and dissimilar input pairs a pushed to have outcome distance greater than the margin m.

$$l_c = \sum_N S * \|W(x_1) - W(x_2)\|^2 + (1 - S) * max(0, m - \|W(x_1) - W(x_2)\|^2) \tag{3.5}$$

The classification loss of the two feed-forward networks are defined as l_{e_1} and l_{e_2} in Equation 3.6 and 3.7 using cross-entropy loss function, where N is the total number of possible pair of inputs.

$$l_{e_1} = -\frac{1}{N}\sum_{i} Ny_1 log(F_1) + (1-y_1)log(1-F_1) \tag{3.6}$$

$$l_{e_2} = -\frac{1}{N}\sum_{i=1} Ny_2 log(F_2) + (1-y_2)log(1-F_2) \tag{3.7}$$

Total loss function L of the designed siamese based network for the regression problem is sum of both feature loss l_c and classification loss l_{e_1} and l_{e_2}. Therefore $L = l_c + l_{e_1} + l_{e_2}$. The ground truths are unit normalized to avoid over-flow in the loss computation.

3.2.3 Forming pairs of samples for regression/classification

The total number of samples in the post-processed data is over 1000 samples. Input pairs are to be formed to train SNN. The total number of possible input pairs is around 0.5M. It is impossible to train the model with 0.5M input pairs, as training process for even one epoch will take impractical amount of time. Hence, a method to select a handful of input pairs just required to train the model effectively needs to be followed.

Input pairs formation procedure:

- To train SNN effectively, the input pairs should not be biased. There should be considerate amount of both similar and dissimilar pairs. Hence, pairs are formed such that it contains equal number of positive and negative pairs. Here, positive pairs are the inputs belonging to the same class and negative pairs are the inputs belonging to different class.

- The input data distribution also needs to be preserved for SNN to be trained effectively. The input data distribution is ensured to be preserved when each sample of input pair is randomly picked during while pair formation.

The input pair formation procedure in SNN classification model is briefed in the algorithm 1. Since, the outcome $conflict$ is the only ground truth that requires a SNN classification model. Therefore, it is a binary classification problem, it has only two classes i.e. class 0 and class 1.

Algorithm 1: Input pairs selection algorithm: SNN Classification

Result: Effectively select the pair of inputs used to train the SNN Classification model
Input: $input\_data, n\_pairs$
Output: $list\_of\_pairs$
Let $input\_data$ be a data-frame contains set of input feature and one of the ground truth;
Let n_pairs be total number of selected input pairs;
Let $list\_of\_pairs$ be set of selected input pairs;
for *i=0 to n_pairs/2* **do**
    Select $class\_x$ randomly from [class 0, class 1];
    Randomly select $first\_sample$ from $class\_x$;
    Randomly select $second\_sample$ from $class\_x$;
    Let $selected\_pair$ contain both $first\_sample$ and $second\_sample$;
    Add $selected\_pair$ into $list\_of\_pairs$;
end
for *i=0 to n_pairs/2* **do**
    Select $class\_x$ randomly from [class 0, class 1];
    Let $class\_y$ be different from that of $class\_x$;
    Randomly select $first\_sample$ from $class\_x$;
    Randomly select $second\_sample$ from $class\_y$;
    Let $selected\_pair$ contain both $first\_sample$ and $second\_sample$;
    Add $selected\_pair$ into $list\_of\_pairs$;
end

The input pair formation procedure needs an additional step to be used for SNN regression problem. In regression, there is no concept of similar and dissimilar class pairs. Hence, this concept is introduced in regression problem by forming classes on the ground truth. Since, the regression outcomes are in the range of 0-100, 10 classes are formed with an interval of 10 each. These classes formed on regression problem is only utilized to form input pairs, whereas the original ground truth is used for training the regression model. The algorithm 2 provides brief procedure on input pair selection for regression model.

Algorithm 2: Input pairs selection algorithm: SNN Regression

Result: Effectively select the pair of inputs used to train the SNN Regression model
Input: $input\_data, n\_pairs$
Output: $list\_of\_pairs$
Let $input\_data$ be a data-frame contains set of input feature and one of the ground truth;
Let n_pairs be total number of selected input pairs;
Let $list\_of\_pairs$ be set of selected input pairs;
Let $classes$ be a list [class 0, class 1, ..., class 10];
for *i=0 to n_pairs/2* **do**
    Select $class\_x$ randomly from [class 0, class 1, ... class 10];
    Randomly select $first\_sample$ from [$class\_x - 1$, $class\_x$, $class\_x + 1$] of $input\_data$;
    Randomly select $second\_sample$ from [$class\_x - 1$, $class\_x$, $class\_x + 1$] of $input\_data$;
    Let $selected\_pair$ contain both $first\_sample$ and $second\_sample$;
    Add $selected\_pair$ into $list\_of\_pairs$;
end
for *i=0 to n_pairs/2* **do**
    Select $class\_x$ randomly from [class 0, class 1, ... class 10];
    Randomly select $first\_sample$ from [$class\_x - 1$, $class\_x$, $class\_x + 1$] of $input\_data$;
    Randomly select $second\_sample$ from $classes$ excluding [$class\_x - 1$, $class\_x$, $class\_x + 1$] of $input\_data$;
    Let $selected\_pair$ contain both $first\_sample$ and $second\_sample$;
    Add $selected\_pair$ into $list\_of\_pairs$;
end

3.2.4 Hyper-parameter tuning & stratified cross-validation

3.2.4.1 Stratified cross-validation

The input data which contains input features and one of the ground truth is divided into 5 folds to perform stratified cross-validation. These folds are formed based on the participant ID i.e. couple ID. Since there exists multiple samples for each participant. It is reasonable to group all the samples belonging to a participant in one fold. There exists a specific behavior pattern in each participant i.e., couple, the behaviour is learnt effectively by the model when all the samples belonging to a couple are present in one fold. Hence, around 8 participant (i.e 8 couples) samples were groups in each fold. Note that there are 39 participants in the post-processed data-set.

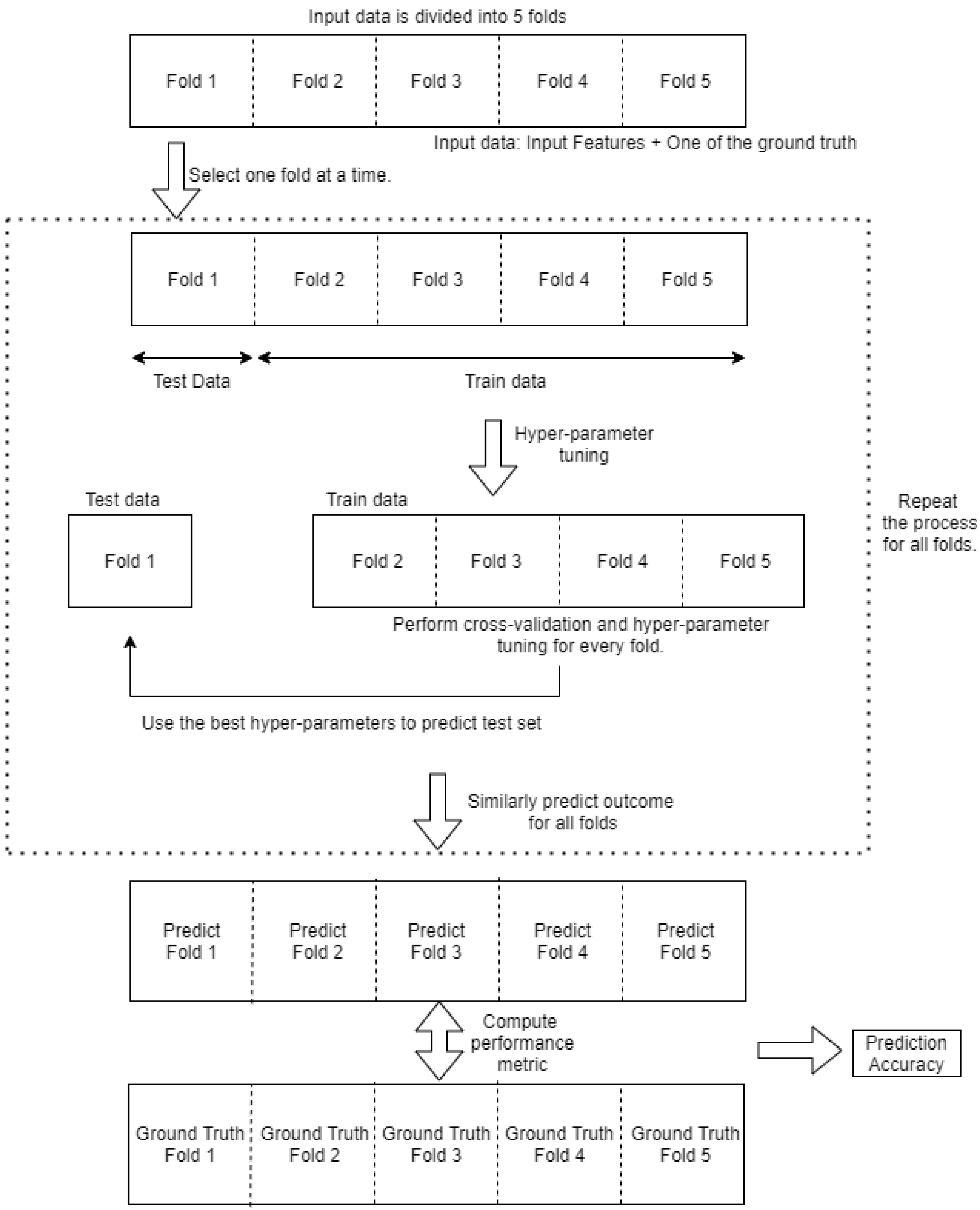


Figure 3.4: Stratified cross-validation

Stratified cross-validation is carried out on the input data as mentioned in the Fig. 3.4. The steps in stratified cross-validation is briefed as follows:

1. Input data is divided into 5 couple-wise folds.

2. One fold is treated as test-data at a time, the rest folds are grouped together as train-data.

3. Hyper-parameter tuning is performed on train-data. Cross-validation on train-data is performed. Here, train-data is again divided into 5 couple-wise folds, each fold predictions are gathered. Performance of the selected hyper-parameter is computed between the gathered predictions for all the folds of train-data and the corresponding ground truth. Finally, the best performing hyper-parameter is chosen for the test-data.

4. The steps 2 and 3 are repeated for all the folds of input-data. The predictions are gathered for the entire input-data.

5. Performance metric is computed using the gathered predictions and ground truth.

3.2.4.2 Hyper-parameter tuning

The table 3.4 provides the hyper-parameter list used in the stratified cross-validation process. In the table 3.4 hidden-layers are the number of hidden layers in the twin structure of SNN, l2-regularization value for all the layers in SNN, dropout values are for all the layers in SNN.

Hyper-parameters	Values
Hidden Layers	1,2
L2 Regularization	0.001, 0.0001
Dropout	0.1, 0.2, 0.3, 0.4, 0.5

Table 3.4: Hyper-parameters

3.3 Methods of personalization

The couples data-set used for the study contains data on multiple participants i.e., couples. The behaviour characteristics of every couple is different. The data-set also includes data on multiple individuals. Therefore, data-set is bound to have high variance. Traditional distribution based models are supervised machine learning models which learn the statistical relation between the inputs and outputs of the data-set. But, they are likely to be ineffective in generalizing the statistical relation when there is a requirement to learn multiple behaviours along with high variance data.

Personalized models are likely to provide a single representation model to learn these multiple behaviours in the data effectively. Therefore, multiple learning models need not be build to extract multiple behaviours from the data. Metric learning procedure provides a possible way to build personalized models. In this work, three variants of personalization has been explored using metric learning via SNN.

3.3.1 SNN for achieving personalization

Personalization has been implemented on the data-set using metric learning method via SNN. SNN inherently ranks the input pairs and provides information on similarity or dissimilarity of the input pairs. Therefore, the model is likely to be capable of learning multiple different behavior characteristics of the data-set.

Figure 3.5 provides a basic personalization model implemented using SNN. The model is trained using the data-set which consists of moment-to-moment multimodal bio-behavior features (i.e., physiological, acoustic and linguistic features) as primary inputs and one of the self-reported daily emotion (i.e., happy, sad, angry, close, nervous and stress) or self-reported conflict as the ground truth. The model learns by minimizing the loss. The loss consists of both metric loss and classification/regression loss. Here, metric loss provides information on how similar or dissimilar the input pair of the model and classification/regression loss provides information on how accurately the classification/regression outcomes were predicted.

The self-reported daily emotions (i.e. happy, sad, angry, close, nervous and stress) have values in the range of 0-100. Hence, the ground truth being any one of the daily emotion is a regression model. The section 3.2.2.1 *Regression* provides details on regression model implementation for personalization via SNN. Feature loss is metric loss for regression model.

The self-reported conflict amongst the couples has values 0 or 1. Hence, the ground truth being couples conflict is a classification model. The section 3.2.2.2 *Classification* provides details on classification model implementation for personalization via SNN. Contrastive loss is metric loss for classification model.

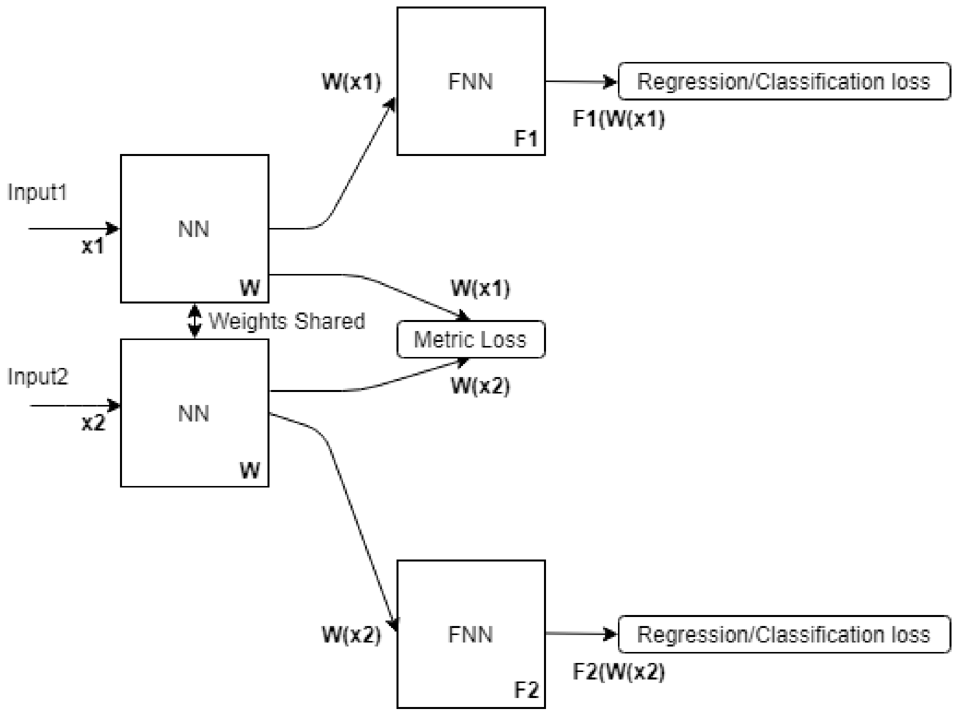


Figure 3.5: Personalization via SNN

3.3.2 Personalization via learning from the same participant (as in the test set)

In the section 3.3.1 **SNN for achieving personalization**, the test-data has participants that are not present in the train-data. If certain samples of participants present in test-data are introduced to the model during training, it can have positive effects on the model performance as there exists chances where the model learns the test-participants behavior. The extent of number of samples of same participants as in the test-data used in model training can also effect model performance.

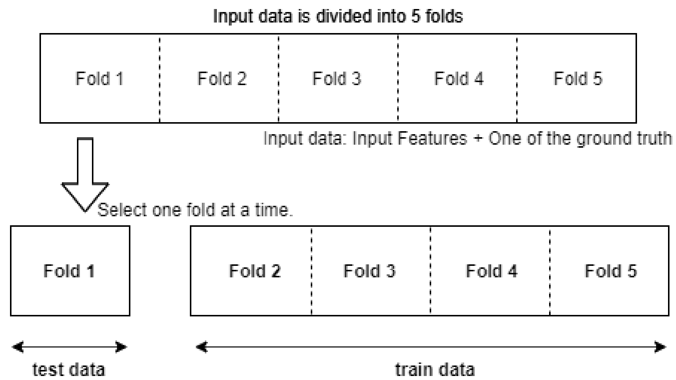
Figure 3.6 provides a detailed procedure on personalization via learning from same participant as test-set in addition to personalization implemented using SNN. Here $fraction$ represents the fraction of samples of participants in the test-data permanently from test-set to train-set. The movement of $fraction$ of samples of every test-participants should be done in such a way that the distribution of ground truth in $fraction$ of samples is same as that of test-data. This movement of $fraction$ of test-data needs to be done prior to cross-validation in the section 3.2.4.1 *Stratified cross-validation*. The sample values of $fraction$ are 20%, 40%, 60% or 80% of samples of every test-participants.

3.3.3 Individuals' relationship characteristic features as additional input to the SNN and in the last layer

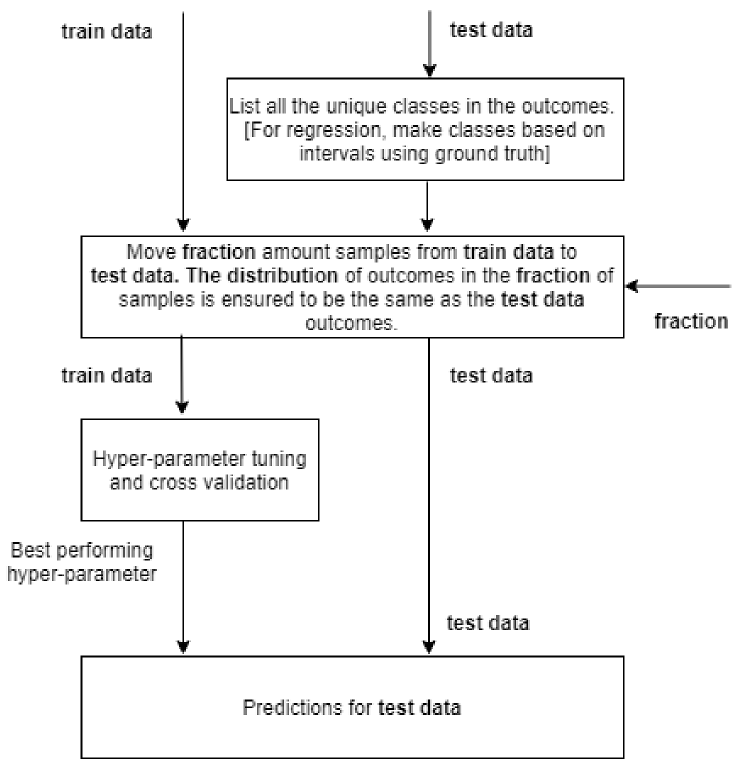
Individuals' relationship characteristic features are obtained from QMI and ECR-R self-assessments. These assessments were provided to the participant couples only once prior to the beginning of the couple study. Relationship characteristics features are used in addition to moment-to-moment multimodal bio-behavior features (i.e physiological, acoustic and linguistic features). These individuals' relationship characteristic features include self-reported relationship satisfaction and attachment data. These additional inputs are used in addition to primary features for achieving personalization. These relationship characteristic features were utilized in two variants of experiments, Figure 3.7 represents the first method where relationship characteristic features are used as an additional input to SNN along with moment-to-moment multimodal bio-behavior inputs and Figure 3.8 represents the second method where relationship characteristic features are added to the last layer.

In both figure 3.7 and figure 3.8, **p1** and **p2** represents the primary inputs i.e multimodal bio-behavior features which include physiological, acoustic and linguistic features. **a1** and **a2** represents the respective relationship characteristic features.

The thought processes behind these experiments were to view the behavior of the model on utilizing the relationship characteristic features prior to and after learning of the moment-to-moment multi-modal bio-behaviour features embeddings using SNN. Relationship characteristics were collected once per participant couple, whereas multi-modal bio-behavioral features were collected hourly though out the day for each participant couple.



(a) Formation of train-data and test-data



(b) Move fraction of test-data to train-data

Figure 3.6: Introduce participant data to model prior to test phase

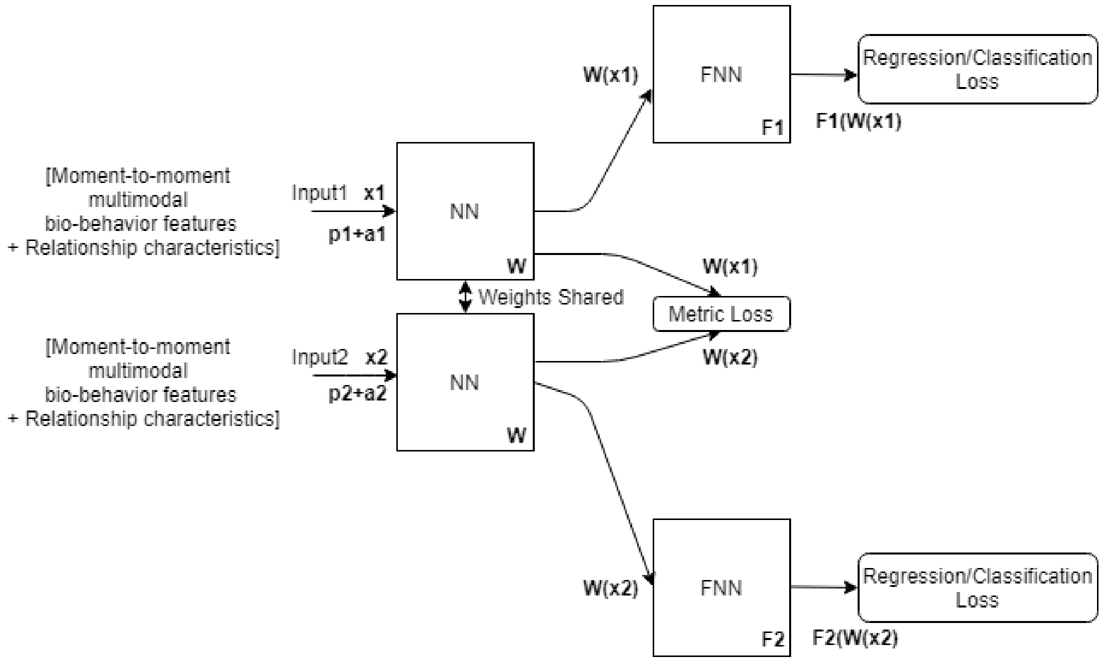


Figure 3.7: Individuals' relationship characteristic features as additional input to SNN

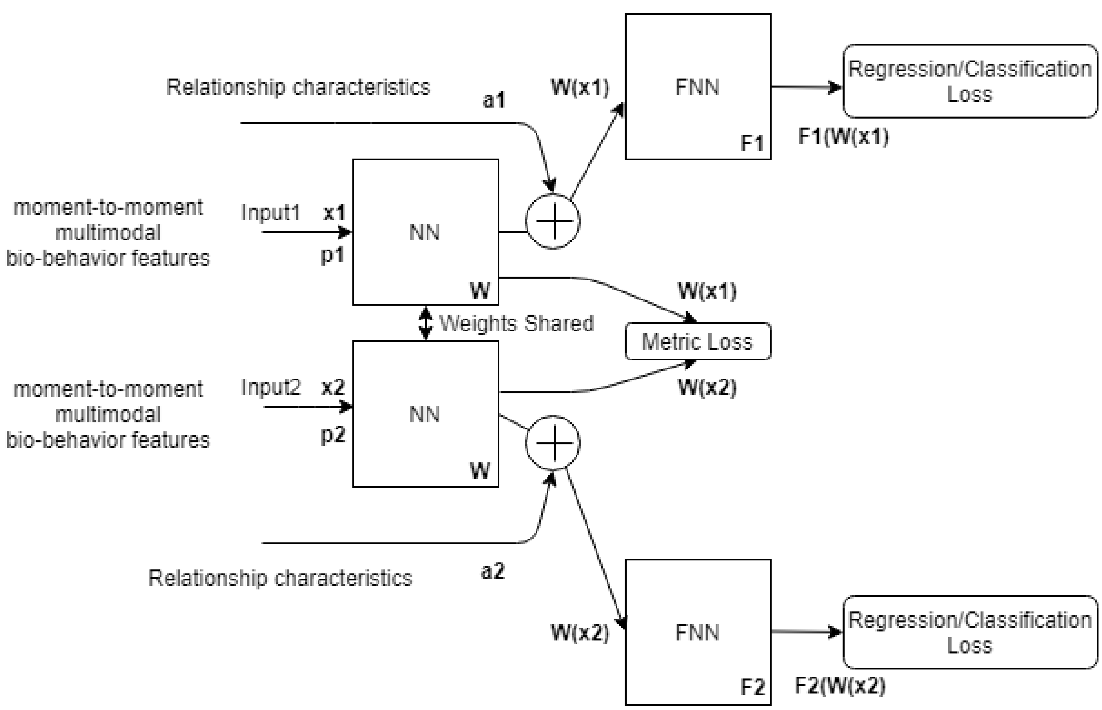


Figure 3.8: Individuals' relationship characteristic features added only to the last layer

3.4 Non-personalized Model

Non-personalized models are basic neural network architecture that is capable of learning the generic distribution of input data and ground truth. Figure 3.9 represents the feed-forward network structure of the non-personalized model used for the classification/regression task. It contains an input layer, two hidden layers and a classification/regression layer. Classification task computes classification loss which is cross-entropy loss and regression task computes regression loss which is mean squared error.

Moment-to-moment multimodal bio-behavior features i.e., physiological, acoustic and linguistic features are primary inputs to the non-personalized model. Individuals' relationship characteristics are used as additional inputs. Self-reported emotions i.e., happy, sad, close, angry, nervous, stress and self-reported conflict are set of ground truths. One of ground truth is selected at a time to perform the experiment.

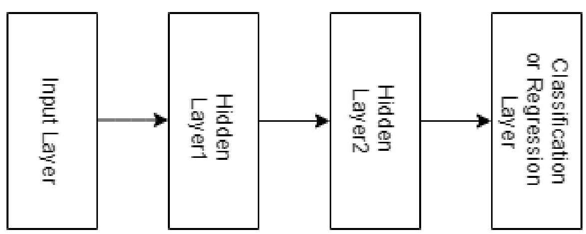


Figure 3.9: Non-personalized Model

4. RESULTS

Personalized models are built using metric learning approach via SNN, methodology details are briefed in the section 3.3. Personalized models follow stratified cross-validation procedure from the section 3.2.4.1. The hyper-parameters used for hyper-parameter tuning are listed in the section 3.2.4.2 which includes number of hidden layers in SNN twin structure, dropout values and l2-regularization for the hidden layers. Hidden layers neurons are 50% of total number of input features respectively. Batch-size is set as 1000 samples. Stochastic gradient descent is used as optimizer. Early stopping is also incorporated. Predicted accuracy were computed using predictions and ground truth. Here, Pearson correlation and p-value used for regression task and f1-score used for classification task. Mean-squared error were computed between predictions and ground truth. Ground-truth for regression model is normalized to avoid NaN's in loss while training, and ground-truth for classification is unvaried from the original data-set.

The following three variants of personalized model classification/regression experiments were conducted:

1. Personalized model with moment-to-moment multimodal bio-behavior features as inputs and one of the daily emotions or conflict as the ground truth. Figure 3.5 briefs the model implementation details. Results are plotted in Figure 4.2.

2. Personalized model with relationship characteristics in addition to moment-to-moment multimodal bio-behavior features as inputs and one of the daily emotions or conflict as the ground truth. Figure 3.7 briefs the model implementation details. Results are plotted in Figure 4.3.

3. Personalized model with moment-to-moment multimodal bio-behavior features as inputs and relationship characteristics as inputs only to last classification/regression layer. Ground truth is one of the daily emotions or conflict. Figure 3.8 briefs the model implementation details. Results are plotted in Figure 4.4.

Non-personalized model implementation details are briefed in Figure **??**. These models also follow the section 3.2.4.1 Stratified cross-validation. The hyper-parameter tuning for these models are performed by tuning dropout values [0.1,0.2,0.3,0.4 and 0.5]. Hidden layers neurons are 50% and 25% of total number of input features respectively. Stochastic gradient descent is used as optimizer. Early stopping is also incorporated. Batch-size is set as 128 samples. Predicted accuracy were computed using prediction outcomes and ground truth. Here, Pearson correlation and p-value for regression task and f1-score for classification task. Mean-squared error were computed between predictions and ground truth. All ground-truth are unvaried from that of original data-set.

The following two variants of non-personalized model classification/regression experiments were conducted:

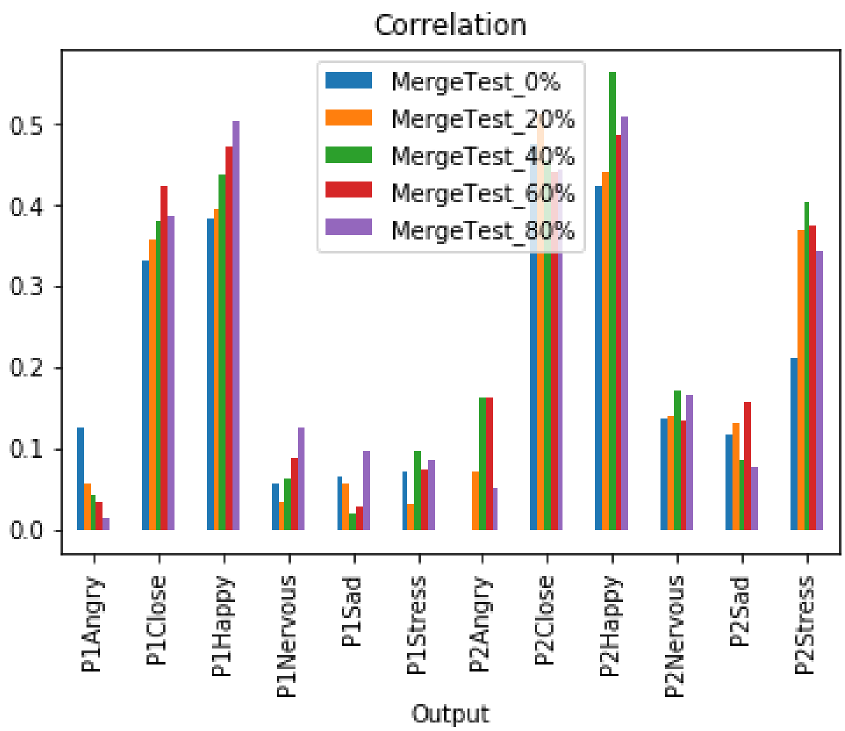
1. Non-personalized models with moment-to-moment multimodal bio-behavior features as inputs and one of the daily emotions or conflict as the ground truth. Results are plotted in Figure 4.6.

2. Non-personalized models with moment-to-moment multimodal bio-behavior features and also relationship characteristics as inputs and one of the daily emotions or conflict as the ground truth. Results are plotted in Figure 4.5.

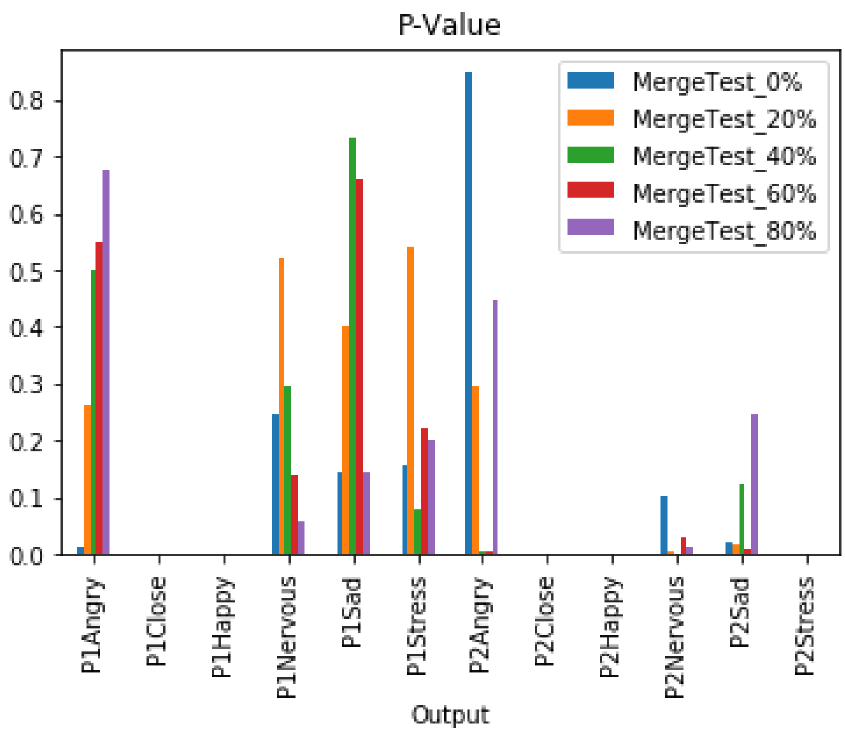
Summary of all the experiments mentioned above is depicted in Figure 4.1. The thesis is on implementation of novel personalized model on USC Couples Data [15] to learn behavioral and well-being characteristics. These personalized models are regression/classification tasks depending on the ground truth. Regression task if ground truth is one of the daily emotions and classification task if ground truth is conflict. The performance of regression model is gauged by gathering Pearson correlation, p-value and mean-squared error values between predicted outcomes and ground truth. The performance of classification model is gauged by gathering f1-score. Personalized model results are compared to that of non-personalized models. Specifically, the results of personalized and non-personalized model with/without relationship characteristics are compared with each other. Note that all the above mentioned experiments results on personalized and non-personalized models includes moving $fraction$ of test-data to train-data as detailed in Figure 3.6. Here, the $fraction$ values are 0%, 20%, 40%, 60% and 80% of test-data. Results on personalized model without relationship characteristics are gathered in figure 4.2, personalized model with relationship characteristics on first layer are gathered in figure 4.3 and personalized model with relationship characteristics on last layer are gathered in figure 4.4. Results on non-personalized model without relationship characteristics are gathered in figure 4.6 and non-personalized model with relationship characteristics are gathered in figure 4.5.

	Common Inputs : Multimodal bio-behavioral features Outputs : One of the daily emotions or conflict	
Model	**Personalized**	**Baseline**
Additional Inputs	No relational characteristics	No relational characteristics
	Relational characteristics on first layer	Relational Characteristics
	Relational characteristics on last layer	

Figure 4.1: Experiments summary

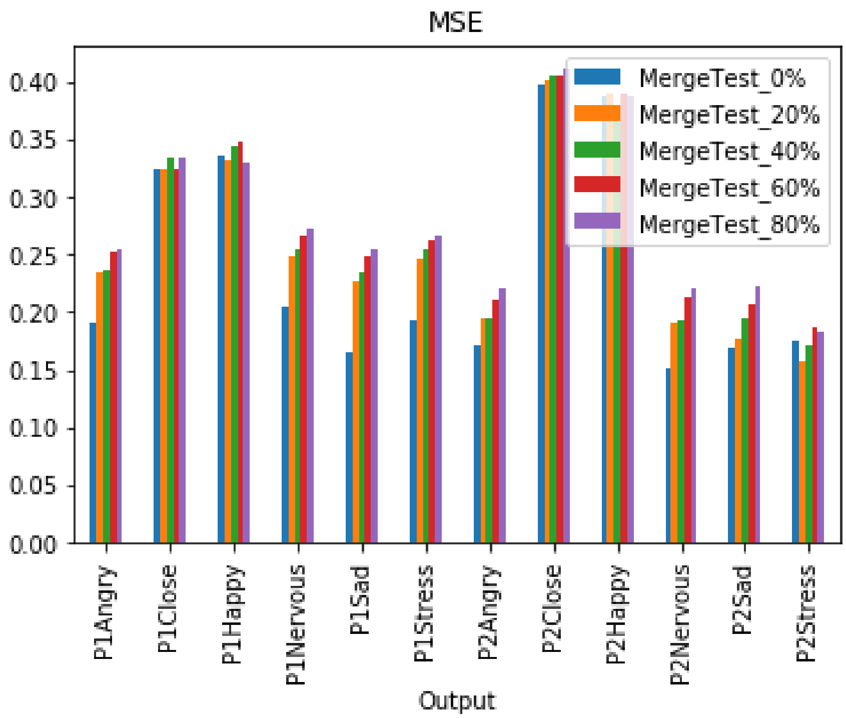


(a) Personalized Models without Relationship Characteristics: Regression Correlation



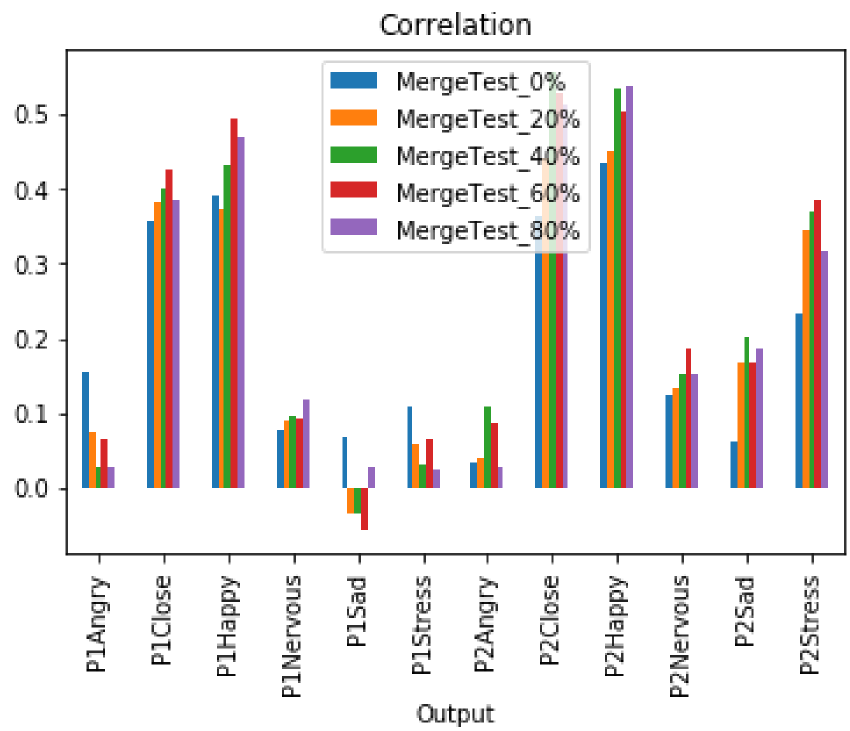
(b) Personalized Models without Relationship Characteristics: Regression P-value

Figure 4.2: Personalized Models without Relationship Characteristics: Regression

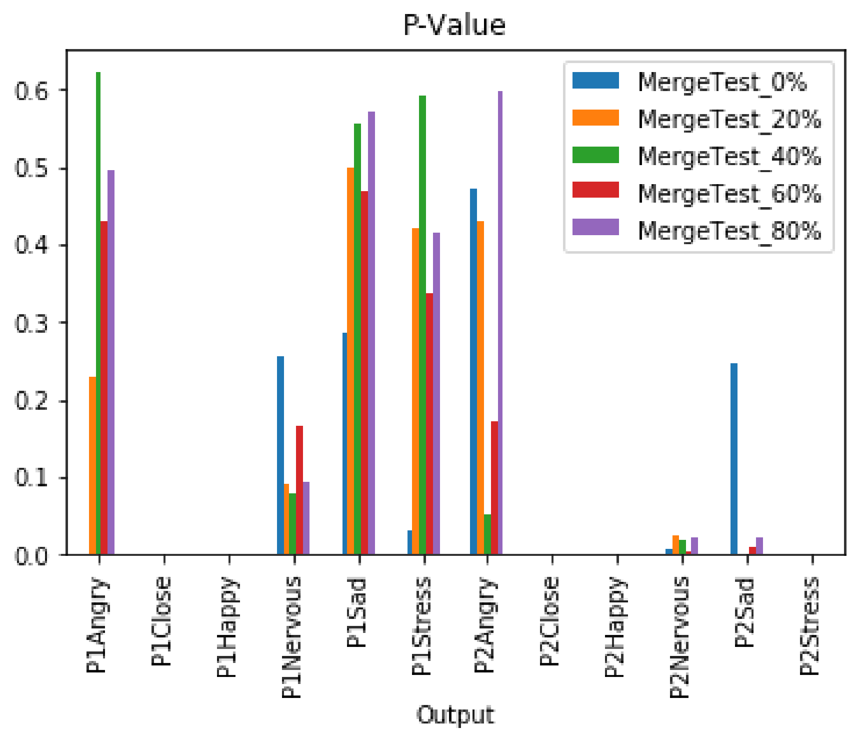


(c) Personalized Models without Relationship Characteristics: Regression MSE

Figure 4.2: Personalized Models without Relationship Characteristics:Regression

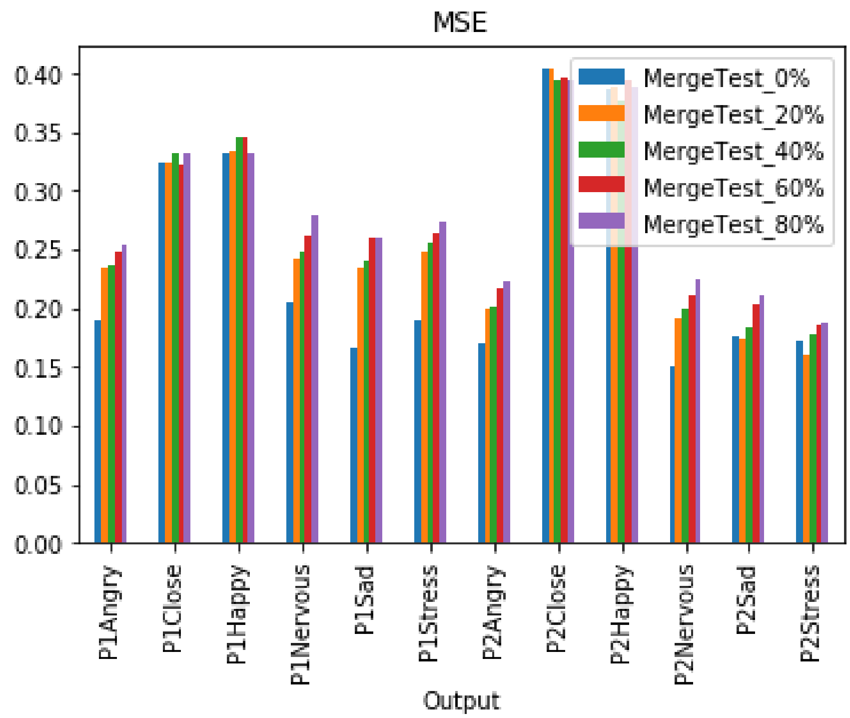


(a) Personalized Models with Relationship Characteristics on first layer: Regression Correlation



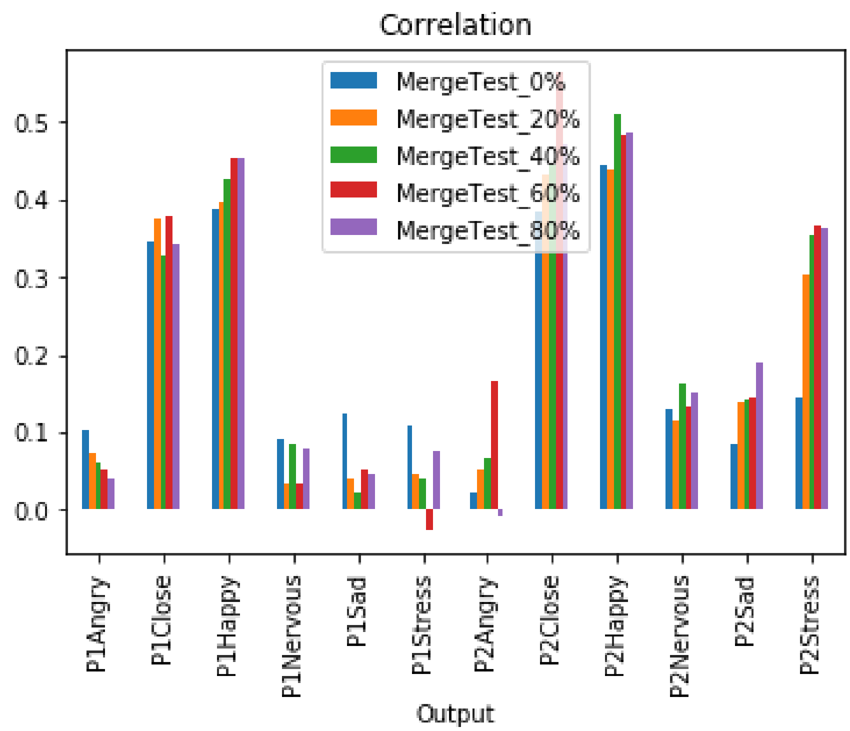
(b) Personalized Models without Relationship Characteristics on first layer: Regression P-value

Figure 4.3: Personalized Models with Relationship Characteristics on first layer: Regression

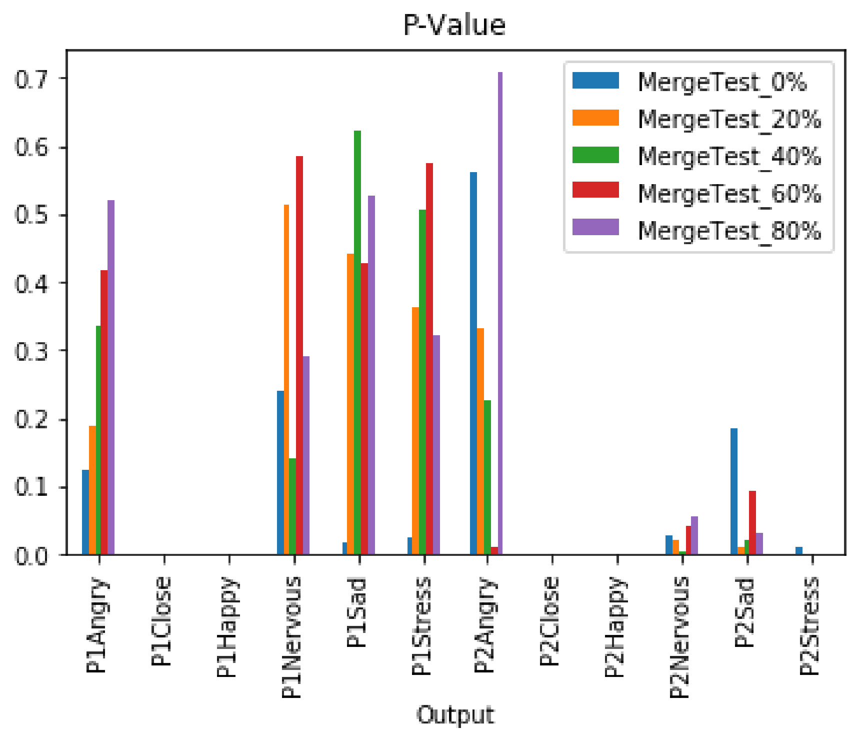


(c) Personalized Models with Relationship Characteristics on first layer: Regression MSE

Figure 4.3: Personalized Models with Relationship Characteristics on first layer: Regression

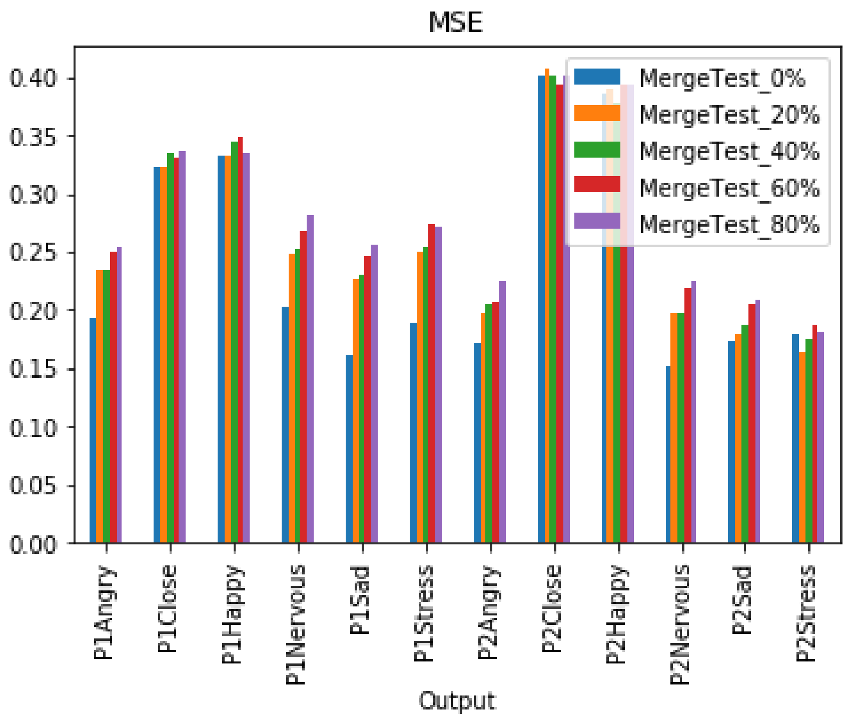


(a) Personalized Models with Relationship Characteristics on last-layer: Regression Correlation



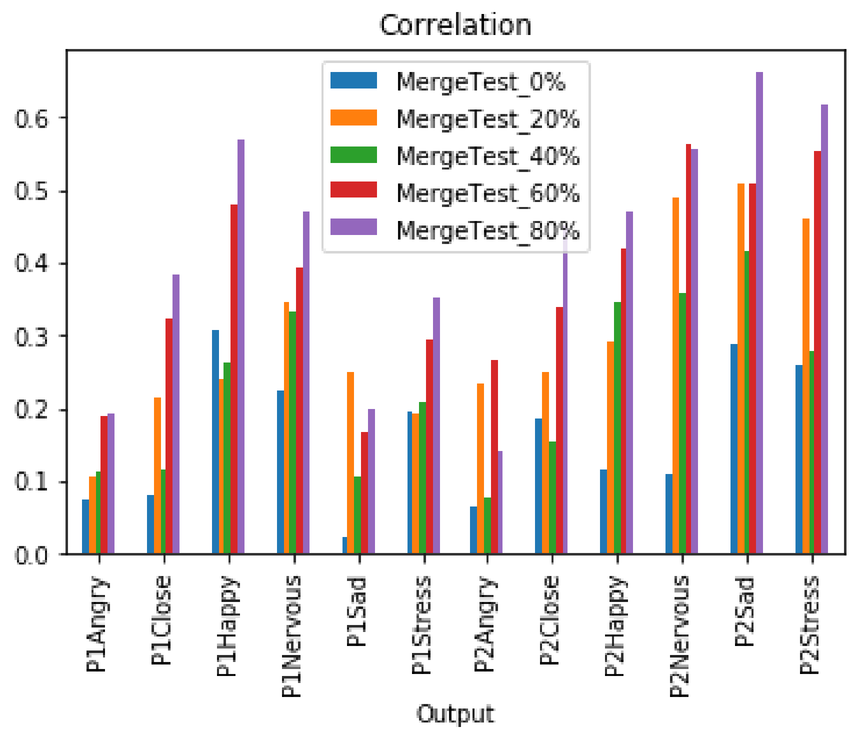
(b) Personalized Models with Relationship Characteristics on last-layer: Regression P-value

Figure 4.4: Personalized Models with Relationship Characteristics on last-layer: Regression

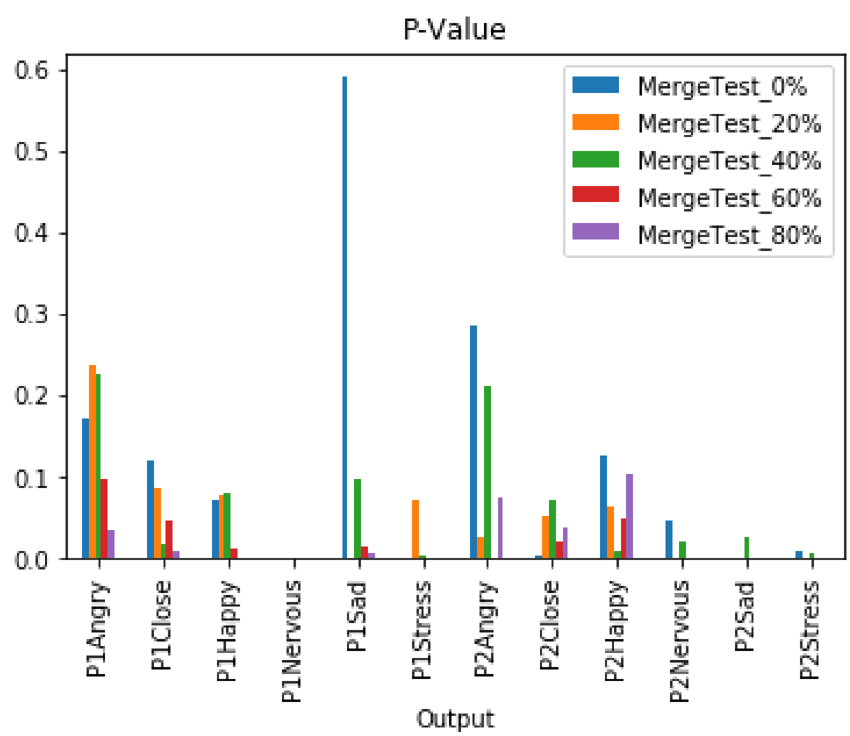


(c) Personalized Models with Relationship Characteristics on last-layer: Regression MSE

Figure 4.4: Personalized Models with Relationship Characteristics on last-layer: Regression

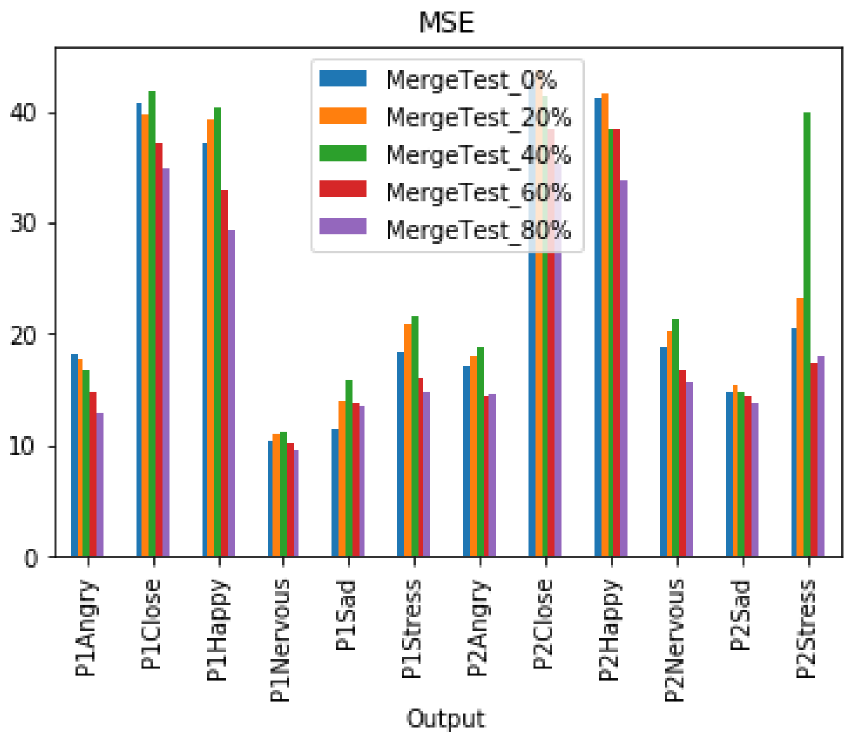


(a) Non-personalized Models with Relationship Characteristics: Regression Correlation



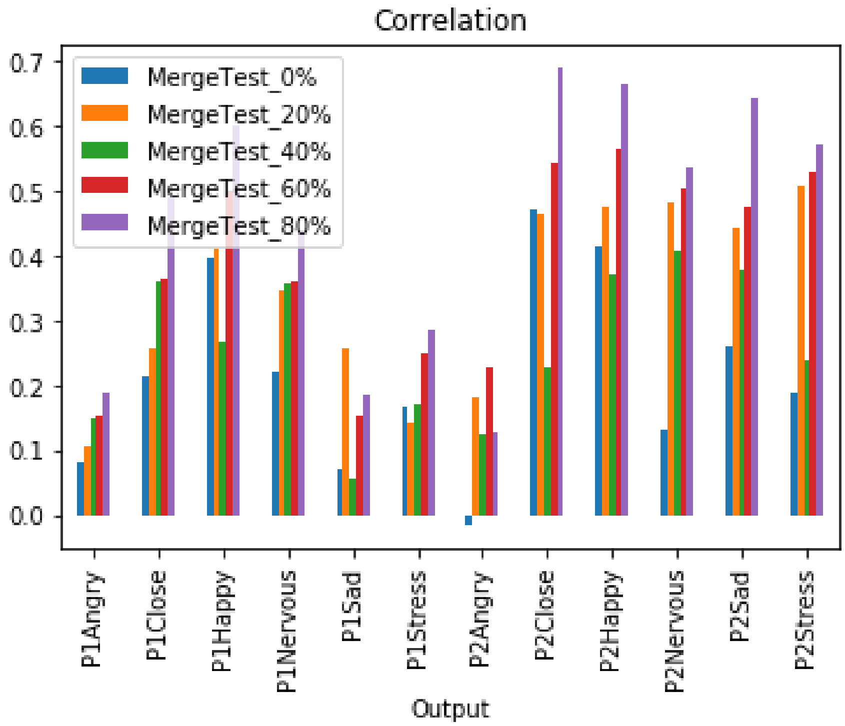
(b) Non-personalized Models with Relationship Characteristics: Regression P-value

Figure 4.5: Non-personalized Models with Relationship Characteristics: Regression

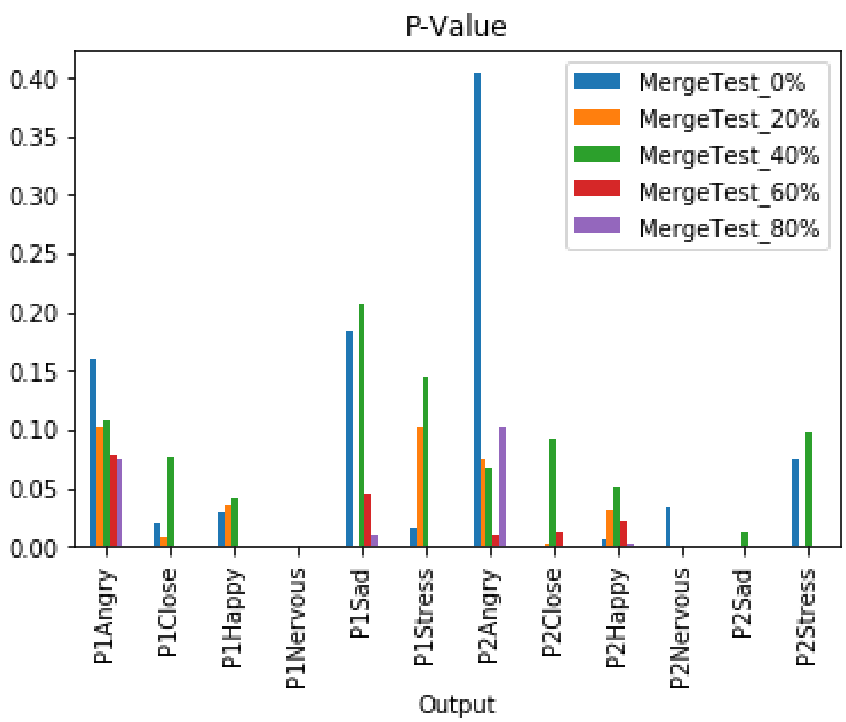


(c) Non-personalized Models with Relationship Characteristics: Regression MSE

Figure 4.5: Non-personalized Models with Relationship Characteristics: Regression

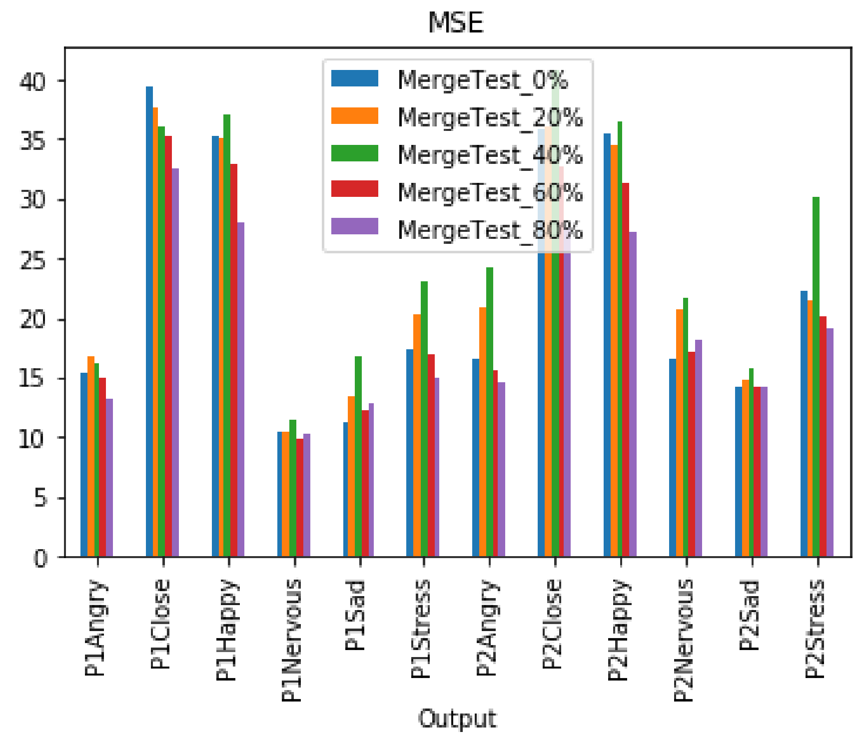


(a) Non-personalized Models without Relationship Characteristics: Regression Correlation



(b) Non-personalized Models without Relationship Characteristics: Regression P-value

Figure 4.6: Non-personalized Models without Relationship Characteristics: Regression



(c) Non-personalized Models without Relationship Characteristics: Regression MSE

Figure 4.6: Non-personalized Models without Relationship Characteristics: Regression

5. DISCUSSION

This M.S. thesis discusses answers for the following aimed research questions:

- **RQ1: To what extent can metric learning approaches learn personalized embeddings of multimodal data for the purpose of detecting emotion-based outcomes and interpersonal conflict?**

  Upon reviewing the result plots, the personalized model implemented using metric learning via SNN does not perform significantly better than the non-personalized models. It is observed that implemented personalized models are extremely sensitive to the pairs of inputs. The possible reason for the depreciated performance of personalized model from that of non-personalized models is likely to be due to the sampled input-pairs of SNN. As, the total number of possible sample input-pairs are reduced from 0.5M to 20000 with the proposed algorithm in section 3.2.3. There is a possibility that the behavioral and well-being patterns of each couple participant are not adequately captured, which could be the possible reason for non-personalized models performing better than personalized models.

- **RQ2: To what extent integrating participants' individual relationship characteristics of participants yields improved performance?**

  The second objective of this work is to integrate partners' relationship characteristics (i.e., satisfaction, attachment) to the model during training. The hypothesis of integrating information about participants' relationship characteristics, in addition to the moment-to-moment multimodal information, did not show any significant improvement in the overall model performance. It is observed that the extracted partners' relationship characteristics are only eight features in number, all are weakly correlated with the ground truth. Hence, adding 8 weakly correlated features to a set of 188 moment-to-moment multimodal bio-behavior features could possibly be the reason for no significant overall performance improved. Figure 5.1 represents the correlation of relationship characteristics with respect to outcomes.

$AnxAtt$ refers to anxious attachment data and $AvdAtt$ refers to avoidance attachment data from ECR-R assessments for both partners whereas $meanQMI$ refers mean of first five questions and $QMI2$ refers to second question of QMI assessment.

- **RQ3: To what extent integrating data samples from a target participant contributes to improved performance?**

  Upon reviewing all the result plots, the hypothesis of utilizing a portion of the target participants' data sample to train the machine learning models, so that the latter can learn personalized patterns explicitly for each individual obtained improved results compared to not including data samples from a target participant in almost all the experiment results.

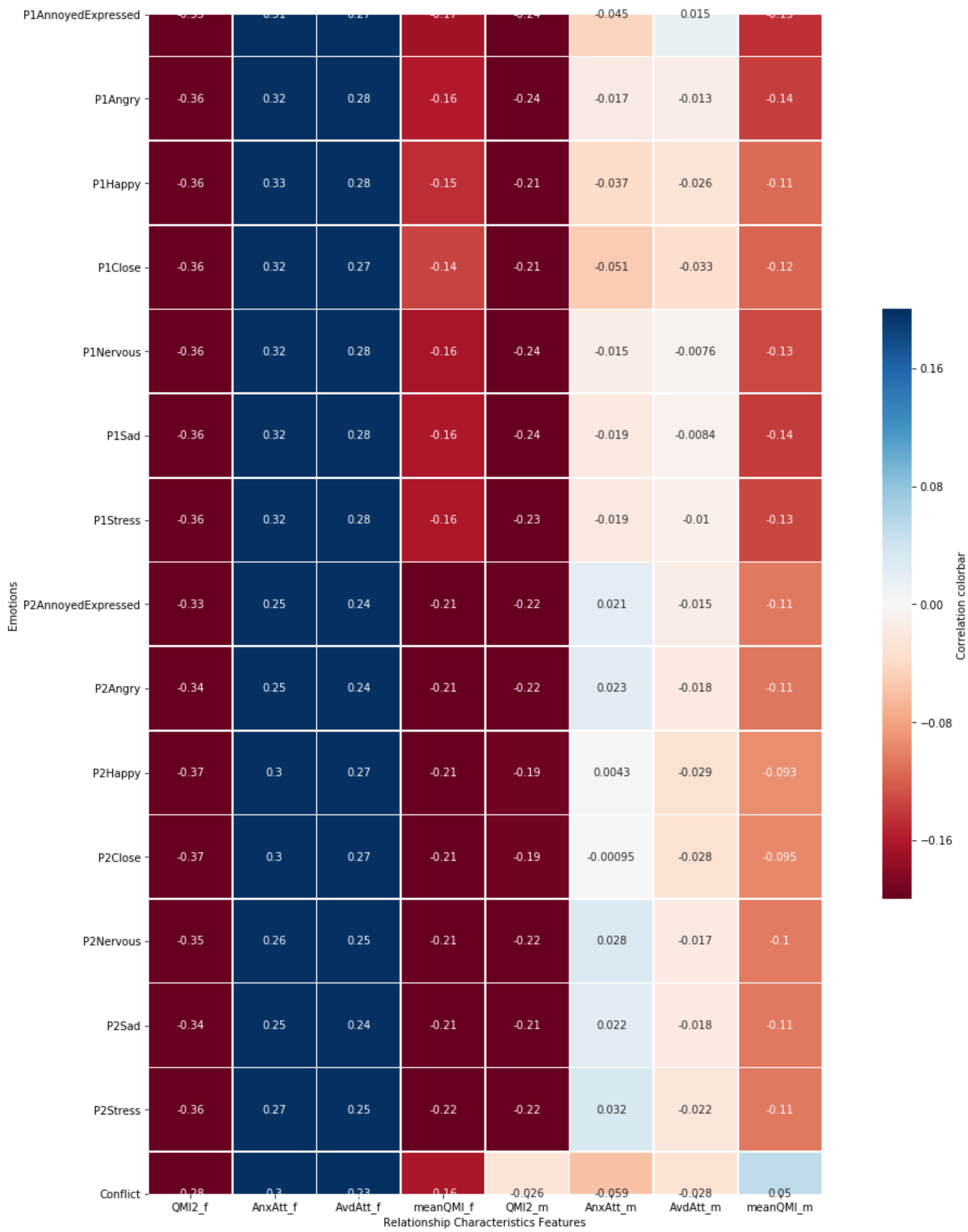


Figure 5.1: Individuals' relationship characteristic features correlation with outcomes

6. SUMMARY, LIMITATIONS AND FUTURE WORK

6.0.1 Summary

A novel personalized machine learning model using metric learning via SNN were implemented utilizing moment-to-moment multimodal bio-behavior signals (i.e., physiological, acoustic and linguistic features) to detect couples daily emotions and detect interpersonal conflict amongst partners. Separate and suitable model architecture were implemented for regression and classification tasks. Individuals relationship characteristics (i.e., satisfaction, attachment etc.) were used in-addition to bio-behavior signals to check if the overall model performance improves. Variants of personalized models were implemented to cater the need to the use of relationship characteristics on the first and last layer of proposed model. Individuals' behavior and well-being characteristics partial introduction during the training phase, to check if there exists any improvement in the overall performance of the proposed model. All the results of personalized models were compared with that of non-personalized model. A non-personalized model is a basic neural network capable of learning generic distribution in the data with respect to the outcome.

6.0.2 Limitations

During the sample pairing process while preparing inputs for the SNN, it is not guaranteed that each participant is equally represented. It is noted that certain ranges of the continuous outcomes are not adequately represented in the data, so it is likely for these behavioral patters to be under-represented during sample pairing process. The study uses ambulatory data collected using smartphones and sensors, therefore they are highly unstructured in nature. Even after utilizing the de-noised data for the study, it is still an inherently challenging task due to the unstructured nature of the data.

6.0.3 Future Work

Revising the sample pairing process such that each participant is adequately represented because the personalized models implemented via SNN are extremely sensitive to input data-pairs.

Since certain ranges of the continuous outcomes are not adequately represented, exploring data augmentation techniques has a possible scope for future study. The ambulatory data used for the study is diverse in terms of demographics. Therefore, examining the integration of demographic features along with multi-modal bio-behavioral features and relationship characteristics features to the proposed personalized model.

REFERENCES

[1] Lumen-Learning, "*Interpersonal Communication*." https://courses.lumenlearning.com/interpersonalcommunicationxmaster/chapter/interpersonal-relationships/. Accessed: 06-10-2021.

[2] M. Mund, B. F. Jeronimus, and F. J. Neyer, "Personality and social relationships: As thick as thieves," in *Personality and disease*, pp. 153–183, Elsevier, 2018.

[3] Wiley-Blackwell, "The way you relate to your partner can affect your long-term mental and physical health, study shows." https://www.sciencedaily.com/releases/2011/06/110617080833.htm. Accessed: 05-07-2021.

[4] A. Brown, "A profile of single americans." https://www.pewresearch.org/social-trends/2020/08/20/a-profile-of-single-americans/. Accessed: 05-07-2021.

[5] S. K. D'Mello, "Chapter 5 - automated mental state detection for mental health care," in *Artificial Intelligence in Behavioral and Mental Health Care* (D. D. Luxton, ed.), pp. 117 – 136, San Diego: Academic Press, 2016.

[6] K. Yan, L. Kou, and D. Zhang, "Learning domain-invariant subspace using domain features and independence maximization," *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 288–299, 2018.

[7] S. Koldijk, M. A. Neerincx, and W. Kraaij, "Detecting work stress in offices by combining unobtrusive sensors," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 227–239, 2018.

[8] K. Gupta, A. Gujral, T. Chaspari, A. C. Timmons, S. Han, Y. Kim, S. Barrett, S. Sichko, and G. Margolin, "Sub-population specific models of couples' conflict," *ACM Trans. Internet Technol.*, vol. 20, Mar. 2020.

[9] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *ICML deep learning workshop*, vol. 2, Lille, 2015.

[10] R. Pflugfelder, "An in-depth analysis of visual tracking with siamese neural networks," *arXiv preprint arXiv:1707.00569*, 2017.

[11] Z. Lian, Y. Li, J. Tao, and J. Huang, "Speech emotion recognition via contrastive loss under siamese networks," in *Proceedings of the Joint Workshop of the 4th Workshop on Affective Social Multimedia Computing and First Multi-Modal Affective Computing of Large-Scale Multimedia Data*, ASMMC-MMAC'18, (New York, NY, USA), p. 21–26, Association for Computing Machinery, 2018.

[12] C. Zhang, W. Liu, H. Ma, and H. Fu, "Siamese neural network based gait recognition for human identification," pp. 2832–2836, 03 2016.

[13] M. Li, K. Chang, B. Bearce, C. Chang, A. Huang, J. Campbell, J. Brown, P. Singh, K. Hoebel, D. Erdoğmuş, S. Ioannidis, W. Palmer, M. Chiang, and J. Kalpathy-Cramer, "Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging," *npj Digital Medicine*, vol. 3, Dec. 2020. Publisher Copyright: © 2020, The Author(s). Copyright: Copyright 2020 Elsevier B.V., All rights reserved.

[14] A. Doumanoglou, V. Balntas, R. Kouskouridas, and T.-K. Kim, "Siamese regression networks with efficient mid-level feature extraction for 3d object pose estimation," *ArXiv*, vol. abs/1607.02257, 2016.

[15] "The usc couple mobile sensing project." `https://homedata.github.io/`.

[16] G. M. Walton, "The new science of wise psychological interventions," *Current Directions in Psychological Science*, vol. 23, no. 1, pp. 73–82, 2014.

[17] H. A. Davis, Mark H. Oathout, "Maintenance of satisfaction in romantic relationships: Empathy and relational competence," *Journal of Personality and Social Psychology*, vol. 53, no. 2, pp. 397–410, 1987.

[18] S. L. Nancy L. Sin, "Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis," *Journal of Clinical Psychology*, vol. 65, no. 6, pp. 467–487, 2009.

[19] R. E. Leslie S. Greenberg, Laura N. Rice, *Facilitating Emotional Change: The Moment-by-Moment Process*. New York: The Guilford Press, 1993.

[20] Substance Abuse and Mental Health Services Administration, Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, *Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health*, 2020. [Online].

[21] S. N. S. e. a. Nahum-Shani Inbal, "Just-in-time adaptive interventions (jitais) in mobile health: Key components and design principles for ongoing health behavior support," *Anals of Behavioral Medicine*, vol. 52, no. 6, pp. 446–462, 2018.

[22] e. a. Stephanie P. Goldstein, Brittney C. Evans, "Return of the jitai: Applying a just-in-time adaptive intervention framework to the development of m-health solutions for addictive behaviors," *International Journal of Behavioral Medicine*, vol. 24, no. 5, pp. 673–684, 2017.

[23] D. S. Thomas, J. Graham Bond, "Behavioral response to a just-in-time adaptive intervention (jitai) to reduce sedentary behavior in obese adults: Implications for jitai optimization," *Health Psychology*, vol. 34, no. Suppl, pp. 1261–1267, 2015.

[24] "Actiwave cardio." `http://www.camntech.com/products/actiwave-cardio/`. Accessed: 02-24-2019.

[25] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1243–1252, 2010.

[26] J. Pennebaker, R. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of liwc2015," 09 2015.

[27] R. Norton, “Measuring marital quality: A critical look at the dependent variable,” *Journal of Marriage and Family*, vol. 45, no. 1, pp. 141–151, 1983.

[28] J. A. Simpson, “Conflict in close relationships: An attachment perspective.,” vol. 71, p. 899, American Psychological Association, 1996.

[29] D. Cramer, “Relationship satisfaction and conflict style in romantic relationships,” *The Journal of Psychology*, vol. 134, no. 3, pp. 337–341, 2000. PMID: 10907711.

[30] D. Cramer, “Facilitativeness, conflict, demand for approval, self-esteem, and satisfaction with romantic relationships,” *The Journal of Psychology*, vol. 137, no. 1, pp. 85–98, 2003. PMID: 12661706.

[31] N. L. Collins, “Working models of attachment: Implications for explanation, emotion and behavior.,” *The Journal of Personality and Social Psychology*, vol. 71, no. 4, pp. 810–832, 1996.

[32] J. Bromley, J. Bentz, L. Bottou, I. Guyon, Y. Lecun, C. Moore, E. Sackinger, and R. Shah, “Signature verification using a "siamese" time delay neural network,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, p. 25, 08 1993.

[33] G. Koch, “Siamese Neural Networks For One-shot Image Recognition,” Master’s thesis, University of Toronto, Toronto, Ontario, 2015.