2023 11th International Conference on Affective Computing and Intelligent Interaction (ACII)

Predicting Loneliness from Subject Self-Report

Liza Jivnani
Dept. of Computer Science
University of South Florida
Tampa, USA
ljivnani@usf.edu

Dr. Fallon Goodman
Dept. of Psychology
George Washington University
Washington, USA
fgoodman@gwu.edu

Jonathan Rottenberg
Dept. of Psychology
University of South Florida
Tampa, USA
rottenberg@usf.edu

Shaun Canavan
Dept. of Computer Science
University of South Florida
Tampa, USA
scanavan@usf.edu

***Abstract*—In this work, we propose to predict four items from the UCLA loneliness scale using subject self-report scores. Using subjective self-reporting, over 14 days, for positive and negative affect, and depression and anxiety we evaluate both subject dependent (personalized) and subject independent experimental design. We evaluate four approaches for prediction, namely random forest, support vector machine, k-nearest neighbor, and logistic regression. We find that the features (self-report) are relatively stable across all four approaches. Along with each individual self-report feature, we also evaluate the fusion of all features where they are concatenated into one feature vector. Through our experimental design, we show that UCLA loneliness items can be predicted, and that the fusion of features (positive and negative affect, and depression and anxiety) is the most encouraging way to do this prediction. We also show that the number of days, used for prediction, has a noticeable impact on the results and that personalization helps with prediction.**

***Index Terms*—loneliness, affect, machine learning**

I. Introduction

Have you ever noticed a kid crying at the entrance of school because they feel like they don't belong? Do they feel lonely at school? "After a long tiring day, I wish I had someone whom I can just talk to!", ever heard anyone saying something similar? Do they feel lonely? Have you ever heard an elderly person say, "I wish I had someone to take care of me?" Do they feel lonely? From a child to an elderly person, anyone can experience loneliness, as up to $2/3$ of Americans suffer from moderate to severe loneliness [1]. Loneliness is defined as a negative experience, that is subjective, resulting from inadequate connections with people that are meaningful [2]. Considering this, there are many reasons why loneliness can occur. For example, it has been shown that low self-esteem is a cause of loneliness [3]. During COVID-19, a study showed that cases of loneliness increased due to the inability to see loved ones from health-related concerns [4]. Interestingly, it has also been shown that there are cases where loneliness scores decreased or did not change during COVID-19, except in cases where the person experienced a de-escalation in a romantic relationship (e.g., divorce) [5]. It has also been shown that children that have trouble communicating are at risk for chronic loneliness [6], as well as children that have been maltreated [7].

Loneliness also has many negative outcomes resulting from experiencing it. Adolescents that report loneliness are at risk for mental and physical health problems, as well as encountering difficulties in education and employment [8]. Multiple studies have shown that social isolation and loneliness can result in all-cause mortality (death by any cause) in older adults [9], [10]. In adults with Autism Spectrum Disorder, it has been shown that increased loneliness can result in anxiety, depression, and suicidal thoughts [11]. Associations have also been found between psychiatric symptoms and loneliness in children [12]. Due to the many negative outcomes from loneliness, it has also been shown that there are many ways that people cope with their loneliness. Consumption is often a way that this is done. For example, consuming products that bring up feelings of nostalgia can restore feelings of belonging and decrease the feeling of loneliness [13]. While social media has been shown to cause loneliness due to both envy and admiration [14], digital technologies have also been used to help people cope with loneliness [15]. People also turn to religious worship or leaders to help cope with loneliness [16]. Other ways to cope with loneliness include music [17], staying active [18], and leisure activities [19].

As loneliness is considered a major public health issue [20], it is important to assess and recognize signs of loneliness [21]. Considering this, we propose to predict four items from the UCLA loneliness scale [22]. More specifically, we train four machine learning classifiers (k-nearest neighbor, logistic regression, random forest, and support vector machine) on subject self-reports of positive affect (PA), negative affect (NA), and depression and anxiety (DEP & ANX). We find that these subjective self-reported features are relatively stable across the four evaluated classifiers, showing encouraging results for each. To evaluate the proposed approach, we developed the following hypotheses to test:

1) **H1:** Loneliness can be predicted using subject self-report over multiple days.
2) **H2:** The number of days, used for prediction, will have an impact on overall predictions.
3) **H3:** Personalization is needed to accurately predict loneliness.

In testing these hypotheses we answer questions such as which set of features (self-report) are best for prediction and do these features generalize across subjects? Overall, the contributions of our work are 3-fold and can be summarized as follows.

1) Subject self-reports are used to predict loneliness on a future day.

2) Subject-dependent vs. independent evaluations are conducted.
3) Impact of the total number of days, used for prediction, is evaluated.

The rest of the paper is organized as follows. Section II gives an overview of related research in the area. Section III-B details the proposed approach, and Sections III and IV detail the experimental design and results, respectively. Finally, Section V gives a discussion on the findings of this work, some limitations, and future work.

II. Related Works

In recent years, research has been conducted on using machine learning for classification and sensing of loneliness. Doryab et al. [23] found that data collected from smartphones and fitbit watches were able to classify levels of loneliness. They showed that these passive sensing technologies are encouraging for college students to detect loneliness and to learn the associated behavior patterns that cause it. Pulekar et al. [24] also used smartphone data that was collected over two weeks. They showed that the big five personality traits [25] are strongly correlated with smartphone loneliness. Using smartphone interaction along with communication features they were able to classify ranges of loneliness. Site et al. [26] used wearble sensors that collected GPS data to learn the relationship between loneliness and the mobility patterns of elderly subjects. They collected indoor and outdoor data and found that the average time spent in these scenarios are important characteristics for analyzing a subject's mobility. Using various machine-learning classifiers (e.g., XGBoost), they found that data collected indoors was better suited for classification, compared to outdoors. Badal et al. [27] used NLP to evaluate sentiment, in interviews from elderly subjects, to classify loneliness. They found that the subjects that were lonely had longer responses with greater expression of sadness.

Along with classifying loneliness, there are other interesting works that investigate other mental health concerns. Asif et al. [28] used EEG signals to classify stress in response to music. They found that different music tracks have a noticeable difference in reduction of stress (e.g., English vs. Urdu tracks). Arif et al. [29] conducted an extensive review on machine learning and classification of anxiety. They found that while there is a lot of work that still needs to be done, this type of classification (e.g., mental health concerns) can help improve health and result in better decision making. Gao et al. [30] used brain scans (MRI) to classify major depressive order (MDD). They found that while results are encouraging, it is difficult to reproduce other results in this area. They conclude that while there are challenges, there is huge potential to use this type of data for MDD. While these works do not directly classify loneliness there is a correlation between these mental health concerns and loneliness [31]. Jannat et al. [32] investigated developmental disorders. More specifically, they showed that self-report along with demographic data can be used to classify autism spectrum disorder. They showed that this data is stable across children, adolescents, and adults. This work motivates our use of self-report data for classification.

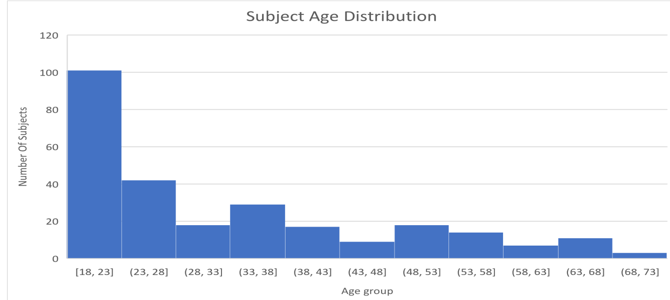


Fig. 1: Age range distribution of evaluated dataset [33].

We extend the state of the art in the following ways: 1) to the best of our knowledge, this is the first work to use subject-self report that includes positive and negative affect, along with depression and anxiety scores for predicting loneliness; 2) we show that when predicting loneliness, the total number of days can have significant impact on the results; and 3) we show that personalized machine-learning has an overall positive impact on results, compared to subject-independent prediction of loneliness.

III. Proposed Approach & Experimental design

A. Dataset

To test our hypothesis that loneliness can be predicted from subject self-report, we made use of the Social Interactions Dataset [33]. This dataset includes loneliness scores and subject self-report. More specifically, it contains 269 subjects (79.2% female, 19.0% male, 1.9% other) with an age range of $[18, 73]$ (See Fig. 1 for age range distribution). The dataset contains measures that include depression and social anxiety symptoms, as well as daily measures that include, but are not limited to, four UCLA loneliness items, time spent socializing, and interaction type. For our experiments we make use of the daily measures including the four UCLA loneliness items, daily measures of positive and negative affect, and depression and anxiety scores, which were each rated on a 5-point Likert scale (e.g., 1=not at all; 5=extremely). Positive affect includes *joyful, relaxed, enthusiastic, and content*. Negative affect includes *irritable, angry, sluggish, and sad*. The loneliness items were drawn from the UCLA loneliness scale [22], and changed to reflect loneliness:

Item 1 (Lone 2): Today, I felt lonely.
Item 2 (Lone 3): Today, I felt that I lacked companionship.
Item 3 (Lone 4): Today, I felt left out.
Item 4 (Lone 5): Today, I felt isolated from others.

1) Pre-processing: We also performed the following pre-processing on this data. 1) If any of the entries in the selected measures were missing, then we discarded data for that day; 2) If a subject participated for less than 3 days, they were excluded. This was done to ensure at least 2 data points in the training data; 3) If the Loneliness score for the selected item didn't change over any of the days, the subject was excluded; and 4) If the output class, we attempt to predict for, doesn't exist in the training set then that data is discarded as well.

B. Proposed Method

As detailed in Section III-A, we evaluated subject self-reported scores from PA, NA, and depression and anxiety. Given this data, we investigate predicting loneliness scores of Nth (N is in the range [3,14] days) day given $N-1$ days of training data. For each feature, we create the following feature vectors $P_A = [d_{i_1}, \ldots, d_{i_4}, \ldots, d_{k_1}, \ldots, d_{k_4}]$, $N_A = [d_{i_1}, \ldots, d_{i_4}, \ldots, d_{k_1}, \ldots, d_{k_4}]$, and $D_A = [d_{i_1}, d_{i_2}, \ldots, d_{k_1}, d_{k_2}]$. Here, d is the self-report for the day, i_j is the first day of training data, and k_j is the last day, where j is the j^{th} feature in the range $[1,5]$. For example, when $N = 5$, $i = 4$, $k = 4$, P_A has a length of 16(4 features $\times 4$ days. Along with these three feature vectors, we also concatenate them together to create the new feature vector $ALL = [P_A, N_A, D_A]$. This is done to test the fusion of all features (Fig. 2). Each of these feature vectors are then used as input to train a machine-learning classifier for predicting the N^{th} day. More specifically, we evaluate the following four classifiers: 1) random forest 2) support vector machine; 3) logistic regression; and 4) k-nearest neighbor.

We treated our experiments as a 5-class classification problem, where separate models were built for each loneliness item. More specifically, there were four models for each experiment we ran, each corresponding to one of the four loneliness items. The output for each model was a prediction for the N^{th} day, where each class was one of the 5-point Likert scale scores for the day. We refer the reader to Section III-A for more details on the UCLA loneliness items and their scale. In total there were 16 (4×4) different models for each experiment - one for each loneliness item (four total) and four feature vectors (P_A, N_A, D_A, and ALL). We use this proposed approach to conduct four different experiments. Three of them are for subject-dependent (personalization), and one for person-independent (generalization). These experiments were done to evaluate if personalization is needed or if a generalized approach toward loneliness prediction is useful. These experiments are detailed in the following sub-sections.

C. Person-Dependent & -Independent Experiments

We conducted person-dependent and person-independent experiments. In total there are three experiments for the person-dependent experiments. Here, each model only contained data from one subject (i.e., training and testing was personalized per subject). We conduct one person-independent experiment (Leave-on-subject-out) to test the generalization of using self-report for predicting loneliness. Experiments $1-3$ below are person-dependent, and experiment 4 is subject-independent.

1) Experiment 1: All-in: We refer to this experiment as *All-in* because all available training data was used for each subject. The experimental goal of the all-in experiment was to effectively predict the loneliness score for the N^{th} day using

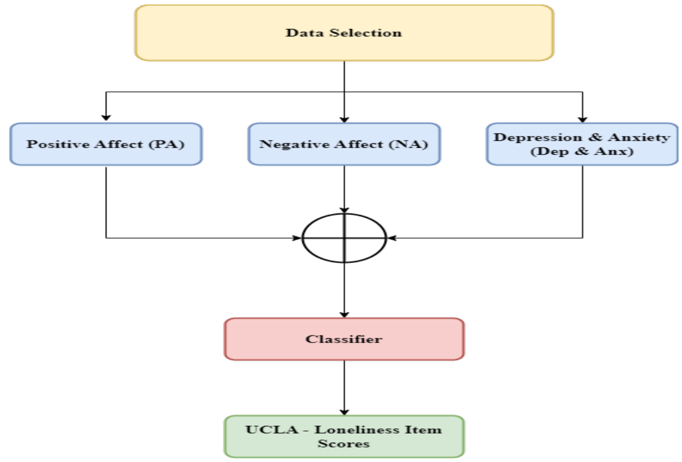


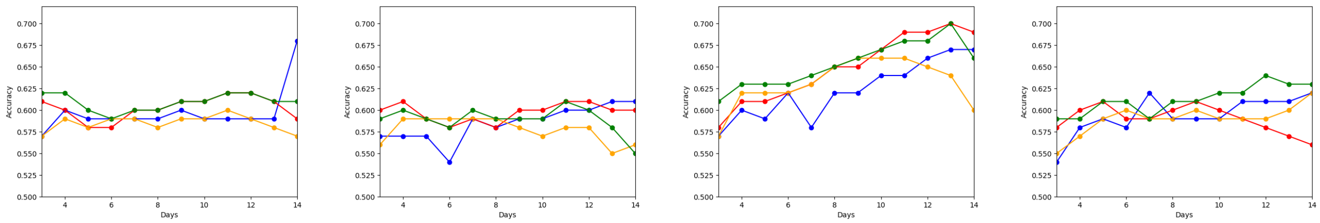
Fig. 2: Overview of experimental design. Data Selection refers to subject dependent vs. independent data. PA, NA, and depression and anxiety features are concatenated and used an input for training. The output is one of five classes - score ($[1,5]$) from UCLA loneliness item. NOTE: This figure shows the concatenation of all features, however, we also evaluate PA, NA, and depression and anxiety features separately.

the data from $N-1$ days where N represents the total number of days of participation for the attempted subject. For example, if a subject participated for all 14 days ($N = 14$), then the training data consists of 13 ($N-1 = 13$) days of self-reported scores, and the 14^{th} day was predicted.

2) Experiment 2: All Combinations: All the possible combinations of N are used, and not just the total number of days that a subject participated. In this experiment, if a subject participated for all 14 days ($N = 14$), then there will be multiple training/prediction runs for each model. More specifically, we tested all possible combinations of N in the range $[3, 14]$. Here, the day that was predicted was not necessarily consecutive. For example (with $N = 14$), we can train on 2 days (e.g., the first two days) of data and test on the 3^{rd} through 14^{th} day. This was done for all possible combinations of days (i.e., training on number of days between $[2, 13]$) and prediction days. While the day that was predicted may not be consecutive, we ensure that the predicted day occurs after the days in the training data.

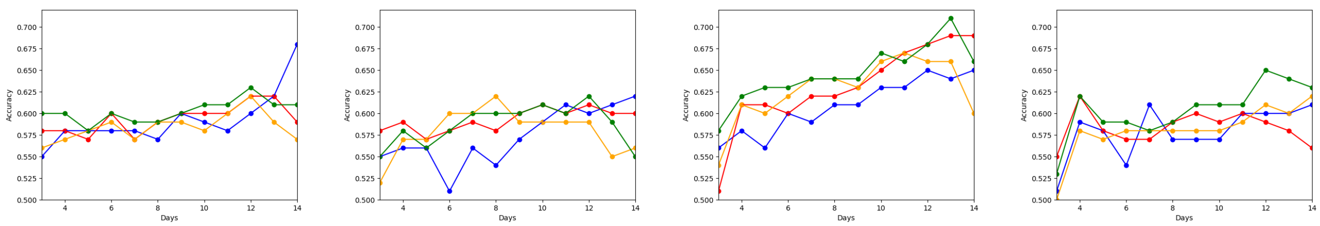
3) Experiment 3: Sliding Window: Here, we used a sliding window, over the days, with a step size of one where the day that was used for prediction was a consecutive day after the training days. For example, again given $N = 14$ we trained on the 1^{st} two days of data, then predicted the 3^{rd}, we then trained on the 2^{nd} two days of data and predicted the 4^{th}. This was also done for $[2, 13]$ days of training data, where the next consecutive day was predicted. The main difference between this experiment and All Combinations is, here, only consecutive days are used for prediction, where All Combinations will also predict non-consecutive days.

4) Experiment 4: Leave-one-subject-out (LOSO): Leave-one-subject-out (LOSO) has the same experimtnal setup as



(a) All Combinations Lone 2 (b) All Combinations Lone 3 (c) All Combinations Lone 4 (d) All Combinations Lone 5

Fig. 3: Average SVM all combinations accuracies across all subjects for predictions of loneliness on each day. Key: Blue: Depression and anxiety; Red: Negative Affect (NA); Orange: Positive Affect (PA); Green: All features (fusion). NOTE: Best viewed in color.



(a) Sliding Window Lone 2 (b) Sliding Window Lone 3 (c) Sliding Window Lone 4 (d) Sliding Window Lone 5

Fig. 4: Average SVM sliding window accuracies across all subjects for predictions of loneliness on each day. Key: Blue: Depression and anxiety; Red: Negative Affect (NA); Orange: Positive Affect (PA); Green: All features (fusion). NOTE: Best viewed in color.

TABLE I: Average evaluation metrics for each experiment for support vector machine. Higher is better for all metrics.

Experiment	Accuracy	F1-micro	F1-macro	MCC
All-in	0.58	0.58	0.48	0.35
All combinations	0.61	0.61	0.53	0.42
Sliding Window	0.6	0.6	0.53	0.41
LOSO	0.48	0.48	0.26	0.06

TABLE II: Average evaluation metrics for each experiment for random forest. Higher is better for all metrics.

Experiment	Accuracy	F1-micro	F1-macro	MCC
All-in	0.56	0.56	0.47	0.33
All combinations	0.6	0.6	0.54	0.41
Sliding Window	0.59	0.59	0.53	0.4
LOSO	0.46	0.46	0.29	0.11

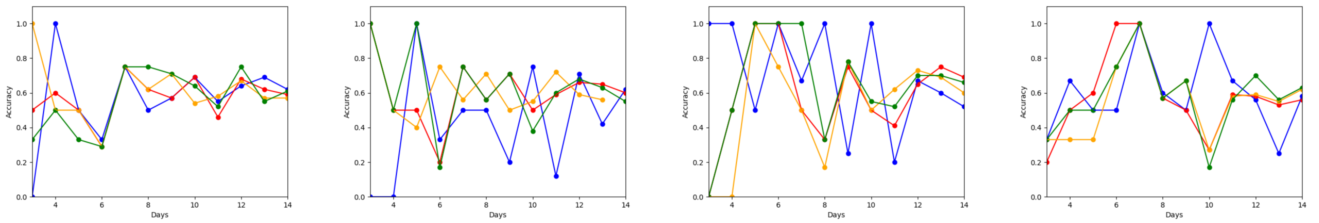
our All-in experiment (experiment 1) in the person-dependent design. We still predict the loneliness score for the N^{th} day using data from the $N-1$ days as training. The difference between the two experiments is that, here, given M total subjects, $M-1$ are used for training data and the M^{th} subject is used for testing. This is done with all subjects used as testing where we ensure the same subject does not appear in both training and testing. It is important to note that we only performed this subject-independent experiment, as there was too much variance in the number of days that subjects reported for. This resulted in the inability to accurately sync up the windows across all subjects for all combinations and sliding window

IV. Results

A. Person-Dependent Experiments (Personalization)

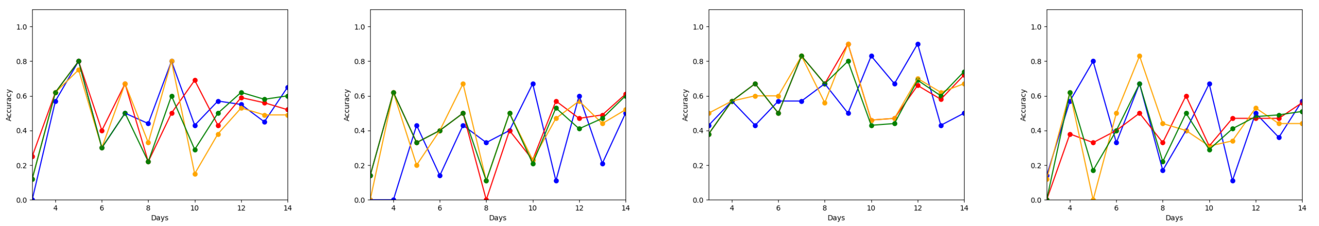
To evaluate our results, we calculated accuracy, F1-micro, F1-macro, and MCC for each of the classifiers. The results for support vector machine (SVM), random forest, KNN, and linear regression can be seen in Tables I-IV. For SVM, as can be seen in Table I, for person-dependent experiments, All Combinations and Sliding Window experiments had similar average results across all metrics. All combinations had 0.61 average accuracy and sliding window had 0.6, same for F1-micro, they both had 0.53 for F1-macro, and 0.42 and 0.41 for MCC. Both of these experiments having similar average results makes sense as they are similar experiments in nature. The main difference is Sliding Window predicts consecutive days, whereas All Combinations allows for prediction of non-consecutive days. Similar trends can be seen for the other classifiers as well (Tables II - IV). The differences in these experiments (e.g., SVM) can be seen when you look at predictions for individual days (Figs. 3 and 4).

From these figures, it can be seen that there are similar trends in the graphs, however, there are some interesting differences in some of the individual days, as well as the different features. For example, for loneliness item 4 (lone 5) it can be seen that there is a larger spike in accuracy



(a) All-in Lone 2 (b) All-in Lone 3 (c) All-in Lone 4 (d) All-in Lone 5

Fig. 5: Average SVM all-in accuracies across all subjects for predictions of loneliness on each day. Key: Blue: Depression and anxiety; Red: Negative Affect (NA); Orange: Positive Affect (PA); Green: All features (fusion). NOTE: Best viewed in color.



(a) LOSO Lone 2 (b) LOSO Lone 3 (c) LOSO Lone 4 (d) LOSO Lone 5

Fig. 6: Average SVM leave-one-subject-out (LOSO) accuracies across all subjects for predictions of loneliness on each day. Key: Blue: Depression and anxiety; Red: Negative Affect (NA); Orange: Positive Affect (PA); Green: All features (fusion). NOTE: Best viewed in color.

TABLE III: Average evaluation metrics for each experiment for k-nearest neighbor. Higher is better for all metrics.

Experiment	Accuracy	F1-micro	F1-macro	MCC
All-in	0.56	0.56	0.47	0.34
All combinations	0.56	0.56	0.51	0.37
Sliding Window	0.55	0.55	0.5	0.36
LOSO	0.41	0.41	0.29	0.09

when using four days of training data compared to three days when the sliding window is used. For the sliding window, with three days of training data, the max accuracy is 0.55 and with four days of training data, the max is 0.625. With all combinations, the max is 0.53 for three days of training data and 0.6 for four days of training data (See Figs. 3d and 4d all combinations and sliding window loneliness item 4). When looking at the individual days, it can also be seen that we are able to predict loneliness with reasonable accuracy. For both all combinations and sliding window, an accuracy of 0.7 and 0.725 were obtained, respectively. This was done using 13 days of training data for UCLA loneliness item 3 (Lone 4 - Figs. 3c and 4c).

For the SVM person-dependent experiment, All-in, evaluation metrics were lower than the average evaluation metrics for combinations and sliding window, however, they are close. All-in achieved 0.58 for accuracy and F1-micro, and 0.48 and 0.35 for F1-macro and MCC, respectively. As can be seen in Table I, this is a difference of < 0.07 across all evaluation metrics. Again, similar trends can be seen in the other classifiers as well (Tables II - IV). While the average evaluation metrics are similar, there is a large difference in the individual days, as can be seen in Fig. 5. This difference occurs between days, and different feature sets. For example, in the all-in experiment, an accuracy of 1.0 was achieved for multiple days and multiple feature types. For loneliness item 1 (Fig. 5a), this accuracy was obtained using the P_A feature vector with two days of training data, and the D_A feature vector using three days of training data. For loneliness item 2 (Fig. 5b), it was obtained using ALL with two and four days of training data. Similar results were found for loneliness items 3 (Fig. 5c) and 4 (Fig. 5d). For loneliness item 3, an accuracy of 1.0 was obtained using training data that ranges from $[2, 7]$ days, using P_A and ALL feature vectors. For loneliness item 4 (lone 5) an accuracy of 1.0 was achieved using five and 6 days of training data, using the N_A feature vector. There is also a larger range between the min and max accuracies for all-in and Leave-One-Sample-Out. For all-in, the min accuracy is 0.0, with two days of training data using the D_A feature vector and the max accuracy is 1.0. The max accuracy was obtained from all of the feature vectors across different days. This contrasts with all combinations and sliding window where the min and max accuracies were $[0.575, 0.7]$ and $[0.5, 0.725]$ for all combinations and sliding window, respectively.

These results support our hypotheses that loneliness can be predicted using subject self-report and that the number of training days used will have an impact on the prediction. More specifically, we have shown that it is possible to predict UCLA loneliness scores on the N^{th} day using using subject self-report training data from $N - 1$ days. Also, by varying

TABLE IV: Average evaluation metrics for each experiment for linear regression. Higher is better for all metrics.

Experiment	Accuracy	F1-micro	F1-macro	MCC
All-in	0.59	0.59	0.49	0.37
All combinations	0.61	0.61	0.54	0.42
Sliding Window	0.6	0.6	0.53	0.42
LOSO	0.47	0.47	0.28	0.09

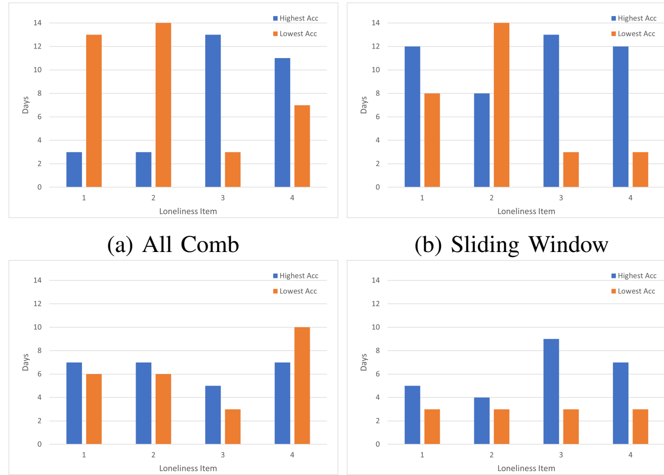
TABLE V: Feature vector with highest and lowest accuracies for each experiment (across all four classifiers) and UCLA Loneliness scale. DEP & ANX=Depression and Anxiety; PA=Positive Affect; NA=Negative Affect; ALL=All Features (PA, NA, and DEP & ANX).

Experiment	UCLA Loneliness	Highest Acc	Lowest ACC
All-In	1	NA	DEP & ANX
	2	ALL	DEP & ANX
	3	DEP & ANX	PA
	4	ALL	NA
All-Comb	1	ALL	PA
	2	ALL	PA
	3	DEP & ANX	PA
	4	ALL	PA
Sliding Window	1	ALL	PA
	2	ALL	PA
	3	DEP & ANX	PA
	4	ALL	PA
LOSO	1	NA	PA
	2	ALL	NA
	3	DEP & ANX	PA
	4	NA	PA

the number of days of training data, we see large variations in the reported evaluation metrics. To further evaluate these hypotheses, we investigate two complementary questions. 1) *Which feature set is best in predicting daily loneliness?* and 2) *Which feature vector is worst in predicting daily loneliness?*

To answer the first question, we summed up the per-model accuracies obtained from each feature vector while predicting each loneliness item. We then evaluated the feature set with the maximum accuracy. For instance: Which of the feature vectors was most accurate in predicting UCLA Loneliness item 1? (Today, I felt lonely). This process was done for each experiment, with results being detailed in Table V (3^{rd} column). Across all the experimental setups, ALL (fusion of P_A, N_A, and D_A) was the most accurate, and D_A was always the most accurate in predicting UCLA Loneliness item 3 (Today, I felt isolated). ALL having some of the highest accuracies across all experiments can be explained, in part, by the fusion of features resulting in improved performance [34]. We also computed the total number of days (N) that resulted in both the lowest and highest accuracies across each experiment. As can be seen in Fig. 7, there is a lot of variance across each experiment. In the majority of models (12 out of 16), when more training data was available, a higher accuracy was achieved. For example, for the sliding window experiment (Fig. 7b), loneliness items 1 and 4 obtained the highest accuracies with 11 days of training data, and item 3 obtained the highest accuracy with 12 days of training data.

To answer the second question (which features are worst), we repeated a similar process, but instead we computed the



(a) All Comb (b) Sliding Window

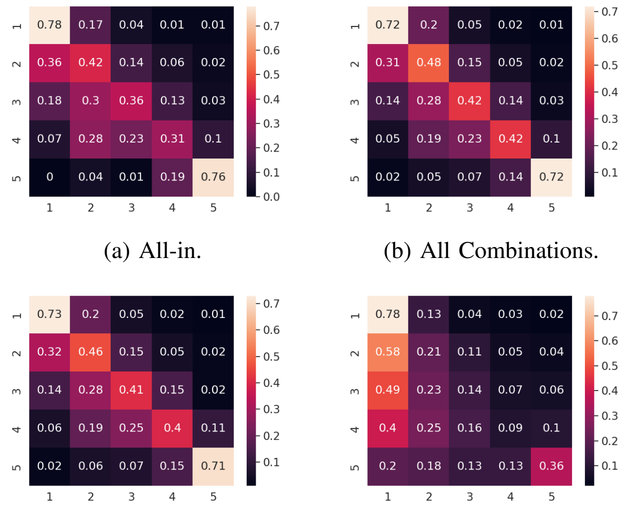
(c) All - In (d) Leave-One-Sample-Out

Fig. 7: Total number of days with lowest and highest accuracies for each UCLA loneliness item, across all classifiers. Key: Blue: Highest Accuracy; Orange: Lowest Accuracy. NOTE: Best viewed in color.

feature vectors with the minimum average sum. As can be seen in Table V (4^{th} column), P_A was often the least accurate feature vector having the lowest accuracies on 12 out of 16 models. The models where P_A were not the least accurate, were all-in for loneliness items 1, 2, and 4, with the D_A feature vector for items 1 and 2, and the N_A feature vector for item 4. We also computed the total number of days that resulted in the lowest accuracies for each experiment (Fig. 7). Overall 4 out of 16 models resulted in a overall lower accuracy when more training data was available. For example, in all combinations (Fig. 7a) the highest accuracy, for items 1 and 2, were obtained with two days of training data. Conversely, the lowest accuracies were obtained with 12 and 13 days of training data for items 1 and 2, respectively. This could be explained, in part, by short-term changes, that can occur, in how lonely people feel [35].

B. Person-Independent Experiments (Generalization)

We have shown that subject self-report can predict loneliness (H1) and that different training days can impact the predictions (H2). We also want to test our third hypothesis (H3) that personalization is needed to accurately predict loneliness. To do this, we conducted a leave-one-subject-out experiment (experiment 4). The average evaluation metrics were much lower compared to the person-dependent experiments. As this experiment was conducted the same as the all-in person-dependent experiment, comparisons with this experiment are most useful. As can be seen in Table I, the average evaluation metrics (SVM) were 0.1, 0.1, 0.22, and 0.29 lower, for LOSO, for accuracy, F1-micro, F1-macro, and MCC respectively. Similar trends, for LOSO, can be seen across the other classifiers as well (Tables II - IV). Also, as seen in Table V, the highest accuracies for each LOSO model had more variation in which feature vector gave the highest accuracy. For example, in the person-dependent experiments, the ALL feature vector



(a) All-in. (b) All Combinations.

(c) Sliding Window. (d) LOSO.

Fig. 8: Average confusion matrices from each experiment, for SVM. The rows/columns correspond to the 5-point Likert scale score for each UCLA item.

resulted in the highest accuracy for loneliness item 4, however, in the LOSO experiment N_A resulted in the highest accuracy. Also, as can be seen in Figs. 6 and 7d, a lower number of days for training data, compared to person-dependent experiments, generally resulted in higher accuracy. For example, loneliness item 1 had the highest accuracy with four days of training data and item 2 had the highest accuracy with three days of training data. Conversely, all-in (Fig. 7c), had the highest accuracies for items 1 and 2 with 6 days of training data for both. These results suggest that there is more variance across subject self-reporting, which makes generalization a much more difficult problem. This also supports *hypothesis H3* that personalization matters when predicting loneliness.

V. Discussion

We presented an approach to predicting UCLA loneliness item scores from self-report. We evaluated a random forest, support vector machine, linear regression, and k-nearest neighbor for prediction. We showed that the self-report features are relatively stable across these classifiers and that our three hypotheses hold true. Namely, H1: loneliness can be predicted using subject self-report; H2: the number of days, used for prediction, will have an impact on the overall prediction; and H3: personalization is needed to accurately predict loneliness. For H3, we further discuss this in Section V-A.

There are also some interesting results that warrant further discussion. Overall, the results show that the Positive Affect (PA) was the least accurate feature set in predicting loneliness. This supports the notion of the negative correlation between Positive Affect and loneliness [36]. Interestingly, despite Negative Affect (NA) being positively correlated with loneliness [36], we find out that in some instances NA was actually the least accurate in predicting loneliness. This occurred in the all-in experiment for loneliness item 4, and the LOSO experiment for item 2. This could be explained, in part, by loneliness being a complex psychological state [37] during which a person feels bad but does not necessarily feel the specific emotions measured in the NA sub-scale (e.g., irritable). This could cause the model to inaccurately predict loneliness.

A. Does Personalization Really Matter?

To further weigh the personalization with generalization, we computed the confusion matrices for each experimental setting (Fig. 8). As can be seen in the figures, the person-dependent experiments show a higher accuracies for each individual item. More specifically, it can also be seen that when the model was incorrect, the item was often predicted as a loneliness item that was close to the ground truth. For example, in the all-in experiment loneliness item 1 was correctly predicted with an accuracy of 0.72, and the next most predicted class for this was item 2 with an accuracy of 0.17. A similar trend can be seen in all person-dependent confusion matrices. Conversely, this trend does not occur in the person-independent experiment (LOSO). As can be seen in Fig. 8d, the majority of items more often predicted item 1 as the correct item. This resulted in items 2-4 predicting item 1 more often than the correct prediction. This could potentially be explained by subjective self-report bias [38], and an imbalance in the data where item 1 was approximately 45% of the classes. This suggests that the person-dependent experiments are able to help mitigate the class imbalance and overall variance in the data. This supports the notion of loneliness being a ”subjective” feeling and the need to take this into account by using subject-dependent modeling [39]. Finally, these results further strengthen our hypothesis H3, suggesting that personalization does matter and is ultimately needed to accurately predict loneliness.

B. Limitations and Future Work

While we are encouraged by this work there are two main limitations of this work. First, only one dataset was used and it is imbalanced. It was imbalanced in several ways: 1) Missing day entries in several participants. Only 30% of all the participants, participated for 14 days. Followed by 16% for 13 days, 12% for 12 days, 11% for 11 days, and 31% distributed almost equally between all the days less than 11 days; and 2) there was class imbalance in the loneliness scores across all items: UCLA score 1 = 45% of the data; UCLA score 2 = 24%; UCLA score 3 = 14%; UCLA score 4= 9%; and UCLA score 5 = 6% of the data. This shows that class 1 dominated the data which is further evident in the person-independent experimental results. This limitation can be potentially addressed in future work using SMOTE [40] to synthesize new data, as well as collecting more balanced datasets. Second, we only used self-report data as our features. Although we have shown encouraging results with this data, other modalities can be used. This includes audio, video, and physiological signals. This can add to the personalization aspect and further strengthen our models and provide some useful insight. Other multidisciplinary factors like personality traits, age, and sex could be used to strengthen the notion of personalization [41], and the overall impact of this work.

Ethical Impact Statement

The work presented in this paper can help people in assessing signs of loneliness, however, there are some ethical considerations with the proposed work that should be taken into account. First, data privacy is a concern. While we are not using modalities that are largely used to identify subjects (e.g., face images), the features are self-reported values of their daily affect. As such, the data needs to be kept safe and all precautions taken to make sure subjects are not identifiable. Second, as we have shown personalization is important for predicting loneliness, this suggests that the prediction does not generalize well to unseen subjects. Considering this, caution should be taken when using the proposed approach on new subjects. Third, while the proposed approach is intended to help people, we need to be careful of any negative applications of the proposed approach. For example, the predictions should be used by a professional to help assess and recognize signs of loneliness, as there can be social stigmas associate with loneliness [42]. Finally, data bias is an ethical concern when using human data. In this case, the imbalance largely came from the number of instance of each class (UCLA loneliness score). With this imbalance, the models could potentially favor scores with more instances (i.e., majority class). This could cause inaccurate predictions, which could lead to unwanted outcomes for those using the approach.

References

[1] "The impact of loneliness on quality of life and patient satisfaction among older, sicker adults," *Gerontology and Geriatric Medicine*, 2015.

[2] L. Fried *et al.*, "A unified approach to loneliness," 2020.

[3] J. Vanhalst *et al.*, "Low self-esteem as a risk factor for loneliness in adolescence: Perceived-but not actual-social acceptance as an underlying mechanism," *Journal of abnormal child psychology*, vol. 41, 2013.

[4] J. N. Beadle *et al.*, "A narrative review of loneliness and brain health in older adults: Implications of covid-19," *Current Behavioral Neuroscience Reports*, vol. 9, no. 3, 2022.

[5] C. D. Ray and S. J. Shebib, "Determinants of loneliness during the covid-19 pandemic in the united states: A one-year follow-up study," *Journal of Social and Personal Relationships*, vol. 39, no. 12, 2022.

[6] C. Crowe *et al.*, "Loneliness prevention and the role of the public health system," *Perspectives in public health*, 2022.

[7] C. de Heer *et al.*, "The association between child maltreatment and loneliness across the lifespan: a systematic review and multilevel meta-analysis," *Child maltreatment*, 2022.

[8] T. Matthews *et al.*, "The developmental course of loneliness in adolescence: Implications for mental health, educational attainment, and psychosocial functioning," *Development and psychopathology*, 2022.

[9] T. L. Barnes *et al.*, "Loneliness, social isolation, and all-cause mortality in a large sample of older adults," *JAH*, vol. 34, no. 6-8, 2022.

[10] B. Yu, A. Steptoe, and Y. Chen, "Social isolation, loneliness, and all-cause mortality: A cohort study of 35,254 chinese older adults," *Journal of the American Geriatrics Society*, vol. 70, no. 6, pp. 1717–1725, 2022.

[11] K. Umagami *et al.*, "Loneliness in autistic adults: A systematic review," *Autism*, vol. 26, no. 8, pp. 2117–2135, 2022.

[12] Y. Xerxa *et al.*, "Childhood loneliness as a specific risk factor for adult psychiatric disorders," *Psychological Medicine*, vol. 53, no. 1, pp. 227–235, 2023.

[13] L. Shrum, E. Fumagalli, and T. M. Lowrey, "Coping with loneliness through consumption," *Journal of Consumer Psychology*, 2022.

[14] Q. Yan, Y. Chen, Y. Jiang, and H. Chen, "Exploring the impact of envy and admiration on social media fatigue: Social media loneliness and anxiety as mediators," *Current Psychology*, pp. 1–14, 2022.

[15] S. Stängle *et al.*, "Use of digital technologies to combat loneliness and social isolation: a cross-sectional study in swiss outpatient care during covid-19 pandemic," *BMC nursing*, vol. 21, no. 1, pp. 1–12, 2022.

[16] J. Gultom and F. Simanjuntak, "Pastoral strategies for the loneliness epidemic of the digital generation," *Jurnal Jaffray*, vol. 20, no. 1, 2022.

[17] S. Cheetu *et al.*, "Understanding the effects of music care on the lived experience of isolation and loneliness in long-term care: A qualitative study," in *Healthcare*, vol. 10, no. 3. MDPI, 2022, p. 457.

[18] S. Marsillas and E. Schoenmakers, "Older adults' mentioned practices for coping with loneliness," *European Journal of Ageing*, vol. 19, no. 3, pp. 753–762, 2022.

[19] S. C. Akhter-Khan *et al.*, "Coping with loneliness in southern myanmar," *Asian Anthropology*, vol. 21, no. 4, pp. 245–262, 2022.

[20] Ł. Mokros *et al.*, "Depression and loneliness may predict work inefficiency among professionally active adults," *International Archives of Occupational and Environmental Health*, vol. 95, no. 8, 2022.

[21] L. W. Holaday *et al.*, "Loneliness, sadness, and feelings of social disconnection in older adults during the covid-19 pandemic," *Journal of the American Geriatrics Society*, vol. 70, no. 2, pp. 329–340, 2022.

[22] "Ucla loneliness scale (version 3): Reliability, validity, and factor structure," *Journal of Personality Assessment*, vol. 66, no. 1, 1996.

[23] A. Doryab *et al.*, "Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data," *JMIR mHealth and uHealth*, vol. 7, no. 7, p. e13209, 2019.

[24] G. Pulekar *et al.*, "Autonomously sensing loneliness and its interactions with personality traits using smartphones," in *HI-POCT*, 2016.

[25] K. Cherry, "What are the big 5 personality traits," *Very well Mind*, 2019.

[26] A. Site *et al.*, "A machine-learning-based analysis of the relationships between loneliness metrics and mobility patterns for elderly," *Sensors*, vol. 22, no. 13, p. 4946, 2022.

[27] V. D. Badal *et al.*, "Prediction of loneliness in older adults using natural language processing: exploring sex differences in speech," *The American Journal of Geriatric Psychiatry*, vol. 29, no. 8, pp. 853–866, 2021.

[28] A. Asif *et al.*, "Human stress classification using eeg signals in response to music tracks," *Computers in biology and medicine*, vol. 107, 2019.

[29] M. Arif *et al.*, "Classification of anxiety disorders using machine learning methods: a literature review," *Insights Biomed Res*, vol. 4, no. 1, pp. 95–110, 2020.

[30] S. Gao, V. D. Calhoun, and J. Sui, "Machine learning in major depression: From classification to treatment outcome prediction," *CNS neuroscience & therapeutics*, vol. 24, no. 11, pp. 1037–1052, 2018.

[31] R. W. Moeller and M. Seehuus, "Loneliness as a mediator for college students' social skills and experiences of depression and anxiety," *Journal of adolescence*, vol. 73, pp. 1–13, 2019.

[32] S. R. Jannat and S. Canavan, "Classification of autism spectrum disorder across age using questionnaire and demographic info," in *ICPRW*, 2021.

[33] G. Silva *et al.*, "What allays loneliness? a fine-grained examination of daily social interactions."

[34] G. Zamzmi *et al.*, "A comprehensive and context-sensitive neonatal pain assessment using computer vision," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 28–45, 2019.

[35] L. Di Blas, M. Borella, and D. Ferrante, "Short-term effects of fluctuations in self-esteem, perceived stress and loneliness on depressive states," *Psihologijske teme*, vol. 30, no. 1, pp. 99–114, 2021.

[36] E. J. Davidson *et al.*, "The relationship between loneliness and positive affect in older adults," *The American Journal of Geriatric Psychiatry*, vol. 30, no. 6, pp. 678–685, 2022.

[37] C. Killeen, "Loneliness: an epidemic in modern society," *Journal of advanced nursing*, vol. 28, no. 4, pp. 762–770, 1998.

[38] A. W. Spithoven *et al.*, "It is all in their mind: A review on information processing bias in lonely individuals," *Clinical Psychology Review*, vol. 58, 2017.

[39] M. H. Lim *et al.*, "Understanding loneliness in the twenty-first century: an update on correlates, risk factors, and potential solutions," *Social psychiatry and psychiatric epidemiology*, vol. 55, pp. 793–810, 2020.

[40] N. V. Chawla *et al.*, "Smote: synthetic minority over-sampling technique," *JAIR*, vol. 16, pp. 321–357, 2002.

[41] D. Barretta, D. Dantzler, and W. Kayson, "Factors related to loneliness," *Psychological Reports*, vol. 76, no. 3, pp. 827–830, 1995.

[42] S. Lau and G. E. Gruen, "The social stigma of loneliness: Effect of target person's and perceiver's sex," *Personality and Social Psychology Bulletin*, vol. 18, no. 2, pp. 182–189, 1992.