PREDICTING BODY DYSMORPHIC DISORDER AMONG DUTCH STUDENTS BASED ON PHONE USAGE USING BINARY CLASSIFICATION MODELS

Student details

Name:	A.J. Rapiej
Student number:	U824818

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN DATA SCIENCE & SOCIETY
DEPARTMENT OF COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

Thesis committee

Supervisor:	dr. A. Hendrickson
Second reader:	dr. H. Brouwer

Tilburg University
School of Humanities & Digital Sciences
Department of Cognitive Science & Artificial Intelligence
Tilburg, The Netherlands
January 2023
Wordcount 8400

Abstract

This research used several binary classification models such as K-Nearest Neighbor, Support Vector Machines, Random Forest, and Logistic Regression to predict Body Dysmorphic Disorder (BDD) based on social media usage. Body Dysmorphic Disorder is a mental disorder that is recognized by disturbing or debilitating concern with minor or imagined flaws in one's physical appearance.

Even though there has been a lot of research predicting mental health disorders, the prediction of BDD has been neglected by many researchers, even though the disorder is very prominent among students. The literature used in this research mainly focuses on predicting depression and anxiety using social media, of which many used different datasets and features. This research distinguishes itself by looking at social media usage and its effect on BDD among students, a topic that has not been researched before.

The best model to predict the body dysmorphic disorder turned out to be a Random Forest with an accuracy of 55.7% using theory driven features. At the end the Random Forest model was tested on disparate groups (gender) and gave a better accuracy score for women group than men group.

Data Source

The data used in this research was obtained from the thesis supervisor, and cannot, and will not be shared outside of the people involved in this thesis. The codes used for this thesis can be found in the link below:

https://github.com/agatarapiej/BDD-prediction

1. Problem Statement and Research Goals

1.1 Context

Over the past years, researcher's interest in social media and the correlation and consequences on student's mental health disorders using Machine Learning models has increased. One of the most neglected mental health disorders by researchers that is prominent among students is Body Dysmorphic Disorder (BDD). As a result, most researchers focus on other mental health disorders in studies of which the findings are vastly different from each other and the influence of these social media applications on a student's body image is unknown. These findings are described in more detail in (See Section 2.4.)

BDD on a person can have serious implications and can interfere with everyday functioning. That is why this research's goal is to discover if BDD on students can be predicted based on social media usage and social media applications to contribute to decreasing the number of students with BDD.

1.2. Research strategy

The research topic and sub-questions are presented in this section. A brief overview of how each question will be investigated is provided under each question. The study topic for examining the influence of social media activity on body dysmorphic disorder and modeling the prediction is as follows:

"To what extent can body dysmorphic disorder be predicted among students, based on social media usage using binary classification algorithms?"

To answer this question, 4 sub questions are formulated:

1) *Which features are important when predicting body dysmorphic disorder?*

The most relevant features for predicting BDD are chosen with Random Forest. The most accurate features are selected based on their highest score. Machine learning algorithms can achieve a greater performance of accuracy in predicting BDD by focusing on the features that are most strongly correlated with the disorder.

2) *Which feature selection technique produces the highest accuracy?*

For this sub question, different feature selection techniques will be analyzed based on the accuracy score of different classification models. As for feature importance, the feature selection can lead to the improvement of the performance of the models by focusing on the most relevant features.

3) *Which classification model predicts body dysmorphic disorder most accurately?*

The performance of different binary classification models is tested on different feature sets, then the performance of each model is evaluated with accuracy and AUC-ROC metric. The highest model performance is selected as most accurate. The reason to identify the best model is that the researchers can acquire a better understanding of the fundamental causes of body dysmorphic disorder and the factors that contribute to their development.

4) *To what extent does the best performing model based on accuracy performs differently when split by gender?*

When developing machine learning models, it is crucial to evaluate how the model will perform on desperate groups. In this research the performance of the model will be evaluated by splitting the test set into separate subsets, men, and women, and measure the model's performance on each subset.

1.3. Findings

The results of the research revealed which binary classification algorithm outperforms the baseline models. Tested on several feature sets, Random Forest scored the highest on accuracy among the models in predicting body dysmorphic disorder. When split by the gender the classification performance of Random Forest achieved higher accuracy for the women group than for the men group.

2. Literature Review

This chapter summarizes significant literature on the research issue. Multiple important phone usage topics are explored in the first part of this theoretical framework. This chapter concludes with a review of previous studies on the prediction of mental health wellbeing.

2.1. Body Dysmorphic Disorder (BDD)

Body Dysmorphic Disorder is a mental health disorder that is recognized by disturbing or debilitating concern with minor or imagined flaws in one's physical appearance (Bjornsson, Didie, & Philips, 2022). BDD has been recognized for more than a century and was discovered by the Italian physician Enrico Morselli in 1891 (Cuzzolaro & Nizzoli, 2018).

According to Veale (2004), Body Dysmorphic Disorder is recognized when a person is obsessed and has a perceived flaw in the individual's appearance. An individual with BDD may spend at least one hour per day engaging in behaviors related to constant obsession with their appearance, such as mirror checking, comparing their body, or seeking cosmetic treatments (Schieber, Kollei, de Zwaan, & Martin , 2015). Even though any bodily area can be the object of worry, the most prevalent preoccupations are with the skin, like scars, acne or color, hair, or nose (Signh & Veale, 2019).

Research of Henn et al. (2019) suggest that in general, more women suffer from a negative body image than men. It is unclear why BDD is more prevalent among women than men. According to Veale et al. (2016) societal pressure and expectations about attractiveness may have an impact. Additionally, negative experience from the past such as bullying and mental abuse or seeing posts on social media that promote an unrealistic idea of beauty may also lead to the development of BDD in woman (Pikoos, Rossell, Tzimas, & Buzwell, 2021). However, it is crucial to remember that BDD may affect both genders, and it is unusual for males to show signs of BDD or seek help by others (Harrison, Fernández de la Cruz, Enander, Radua, & Mataix-Cols, 2016).

2.2. Social Media Influence

Certain social media platforms may have a greater impact on appearance-related disorder than others. Platforms such as Instagram, Facebook and Snapchat place a great focus on self-presentation and confirmation-seeking activities which can lead feeling under pressure to show an ideal version of themselves publicly to others (Sharma, John, & Sahu, 2020). According to Fardouly and Vartanian (2016) social media use is connected in several ways with higher body dissatisfaction and low self-esteem, especially among young adults and students, who are more sensitive to societal and media influence. Due to big user base of this platforms, users are exposed to a diversity of pictures of bodies and faces, that fall far away from reality affecting the view and satisfaction with own appearance (Karsay, Trekels, Eggermont, & Vandenbosch, 2021). Furthermore, platforms such as Instagram allow users to interact with each other and develop communities, which may be helpful, but it can also have a negative side of it as create a feeling of constant comparison and competitiveness with the peers (Carey, Donaghue, & Broderick, 2014).

2.3. Feature selection techniques

To support this research, multiple similar papers researching mental health disorders were investigated that used different feature selection techniques. These papers may give a glance into what feature selection models could be effective to use in a study that is mental health related, i.e. BDD. The range of feature selection techniques used in these studies are different per research, and can be used to support this study's methodology approach.

Zulfiker et al. (2021) researched predicting depression based on the Burns Depression Checklist (BDC) data. They applied three different feature selection techniques such as ANOVA f-test, minimum redundancy and maximum relevance (mRMR), and the Boruta feature selection algorithm. The research claims that decreasing the number of features with feature selection techniques not only has increased the accuracy score of the model but also reduced the running time of the models. Furthermore, it was discovered that ANOVA feature selection method surpassed other techniques and achieved the highest classification accuracy of 92.56% with the Adaboost classifier.

Where Zulfikier et al. (2021) only focused on the BDC survey, the research of Angskun et al. (2022) combined Twitter user's information and a self-reported survey from … to predict depression. To eliminate unnecessary features, the Random Forest, ANOVA and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) were used. The findings demonstrated that ANOVA achieved 76.57% accuracy on Adaboost and outperformed SVM-RFE (76.46%).

The research of Zulfiker et al. (2021) and Angskun et al. (2022) used a combination of feature selection techniques, however, Ernala et al. (2019) eliminated associated features and increased the model predictability by reducing the noisy features with only the ANOVA F-test. With this feature selection method, the features were reduced from 550 to k-best features per model. The study was focused on using social media data to forecast people's mental health statuses. The features contained the linguistic information from Facebook and Twitter accounts from the participants activity such as comments and shared posts. The highest model performance achieved 75% accuracy having the lowest number of features (350) compared to other models.

2.4. Mental Health Prediction Models

Social media data may be beneficial for researchers and healthcare experts since it allows them to obtain important insight into people's thoughts and activity. The data might be used to discover different patterns that can suggest the absence of a mental heal issue as well as tracking a person's behavior online.

Almouzini et al. (2019) collected Arabic tweets from 89 Twitter user to predict whether a tweet represents any signs of depression or not. The data was trained on 4 different machine learning algorithm such as Random Forest, AdaBoost, Liblinear and Naïve Bayes. The Random Forest outperformed other classifiers in terms of accuracy (73.5%) followed by Liblinear classifier (71.2%) and AdaBoost obtained the lowest accuracy of 55.2%.

Bakar et al. (2021) tried to detect depression level from tweet data from more than 10,314 random users that was obtained from a scrapped tool. Focusing on more linguistic approach. The research claims that when training the model on the tweet data, the Neural

Network accuracy and recall results are exactly the same as those of SVM, 78.27% and 0.042 respectively. The third machine learning model used in this research, Decision Tree did not perform well, this demonstrates that the Neural Network and SVM models are the most accurate in predicting depression level using large dataset from social media.

Gkotsis et al. (2017) gathered data from different social media platform such as Reddit, where the posts from 32,280 users were analyzed, and classifiers were built to detect and categorize posts concerning mental illness based on 11 disease themes such as depression, self-harming etc. The convolutional neural network (CNN) identified the relevant topics and recognized mental illness-related posts from balanced dataset with an accuracy of 91.08%. The Feed Forward (FN) came in second with 90.78%, followed by SVM 85.87% making the Linear Regression the last with 84.01% .

Angskun et al.(2022) build a depression detection model feeding it with both demographic factors and tweets as input features and self-reported Patient Health Questionnaire-9 answers as an outcome feature from around 230 participants. The study investigated five machine learning techniques Support Vector Machine, Decision Tree, Naïve Bayes, Random Forest, and Deep Learning. The findings demonstrate that the Random Forest model detected depression with 74.35 % accuracy, achieving the highest performance among other models.

2.5. Summary

The majority of research on predicting mental health disorders among social media users focused on user-generated content. This involves the content of the tweets on Twitter, status updates on Facebook, and pictures shared on Instagram. Other information generated by an individual are activities captured by phone logs, how much time an individual spends on a specific application, and the number of sessions that are not covered in any of the research. This research uses a more simplified dataset, where specific user generated content is not available. As a result, the most accurate feature selection methods and mental health prediction models described in the literature are used as inspiration, but it cannot be expected for them to have the same accuracy in this research. Unfortunately, even though there are many researchers investigating multiple mental health disorders in relation to social media usage, there has not been any research that has tried to find a correlation between BDD and social media usage. To bridge that gap, this study focuses on identifying which models can best predict BDD, in relation to social media usage.

3. Methodology

In this research, a systematic and organized strategy is used for preprocessing the data, use of adequate feature selection models, hyperparameter tuning, evaluation of machine learning models performance and selection of most accurate model for predicting the Body Dysmorphic Disorder (Figure 1) .

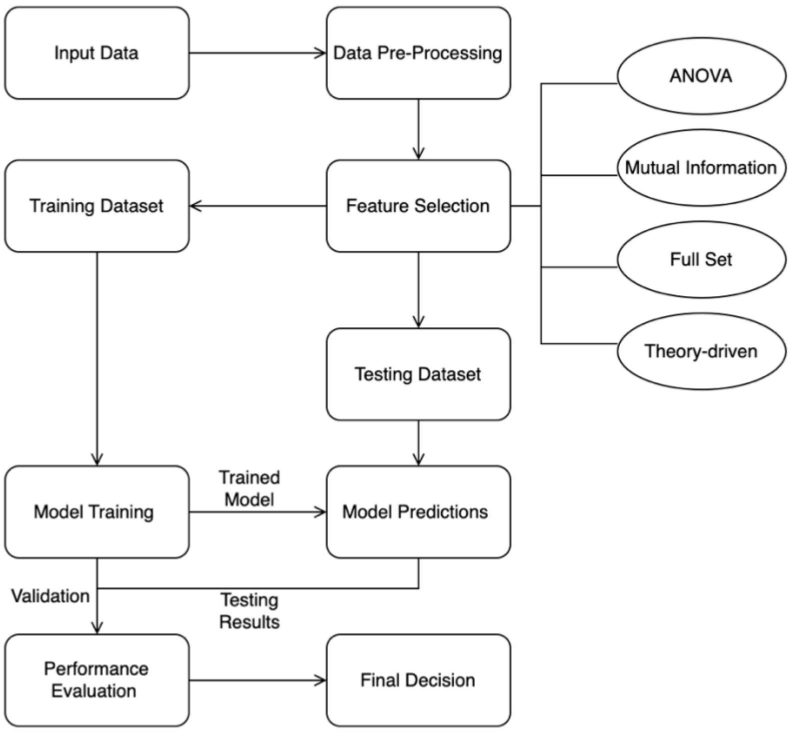


Figure 1. A methodology framework for ML prediction of BDD

3.1. Dataset description

The dataset for this research is obtained from a study done by George Aalbers, Andrew T. Hendrickson, Mariek M. P. vanden Abeele, and Loes Keijsers. The dataset was extracted upon investigating how smartphone usage and mental health are statistically related. The dataset is made up of application logs captured by the Ethica Data v151, health research software platform, from approximately 240 Tilburg University students between February and June 2020.

3.1.1. Application dataset

The application events dataset contains a log of continuous data collected from application usage. The dataset is made up of round 240 application events csv files that provide information about 11 variables. Each row provides information from a single session. The variables contain information about the phone user's id, the application url, and the time they opened and closed certain application. Moreover, the smartphone’s battery level is logged, as well as the session id and a binary variable providing information on whether the user got a notice before accessing the application.

3.1.2. Mood dataset

Several mental health related questions were used to assess the psychological well-being level of each participant during the survey. This study used variety of questions related to mental health being and divided it into 3 categories. The first 21 questions were about the depression symptoms, the next 15 questions were related to burnout. Finally the last 11 questions inquired body dysmorphic disorder. Participants had to rate their feelings on a Likert scale of 1 to 5. The following were the potential responses on a 5-point Likert scale: “None”, “Mild”, “Moderate”, “Severe” and “Extreme”. Additionally, per survey the following demographic characteristics such as the user ID, age, sex, gender, phone model and data version were collected.

3.1.3. Application categories dataset

The application categories dataset included all information on the applications collected in the participants' phone logs, such as application id, count, category, and application name.

3.2. Preprocessing the data

Data processing of the mood data primarily contained of extracting relevant features for the research. Each data column was examined independently to see if any of the values were erroneous. Investigation was accomplished by utilizing boxplots to determine whether or not the outliers were present. Figure 2 represents boxplot of a variable age. It can be seen that the one of the participants when answering the question about the age answered it with a zero value, which is not possible, hence this value was removed to create a robust model (Sarker, 2019). Furthermore, when looking at the gender, 3 unique values were recognized: female, male and others. Two users identified itself with other value and were assigned to the majority class which is female.

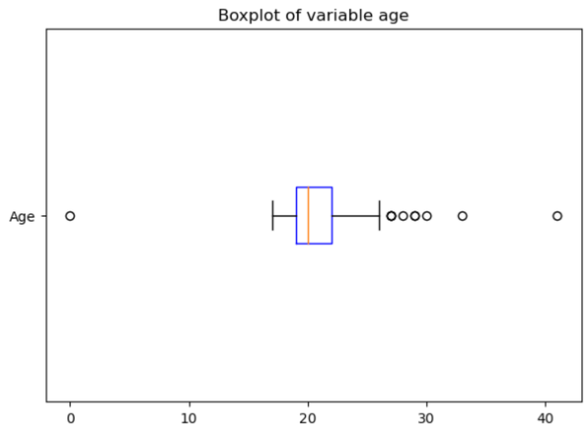


Figure 2. Boxplot of variable age

To improve the performance of machine learning models the raw data was transformed into new features that can be used in predicting body dysmorphic disorder based on the social media data. First, the time spent in minutes was computed from starting time and end time and the number of total sessions was computed from all sessions that are present in the data. Additionally to analyze the activity of the participants during the day, the day of the week variable and daytime variable 6 daytime window of 4 hours was created: late-night, early morning, morning, noon, evening, and night.

The variables start time, end time, notification, notification Id, battery, and survey Id were deleted after generating new variables since they did not provide any value to the research. To improve the prediction performance of the data, values with 0 time spent were removed. Furthermore, the median, maximum, and minimum amount of time spent on social media applications was computed.

3.2.1. Predictor variable

The focus of this study is on the social media category therefore nonsocial category data was not considered for creating social media variables. Table 1 represents the final features used in this study, all of them are continuous variables.

Table 1. Final features

Variable	Description	Measure
total_sessions	Total number of sessions	Count
time_sessions	Total time of sessions	Sum
sns_mean_time	Mean time of social media application	Mean
sns_max_time	Maximum time spent on social media	Max
insta_time	Total time spent on Instagram sessions	Sum
insta_sessions	Number of Instagram sessions	Count
fb_time	Time spent on Facebook	Sum
fb_sessions	Number of Facebook sessions	Count
snap_time	Time spent on snapchat	Sum
sns_sessions	Total number of social media sessions	Count
sns_time	Total time spent on social media applications	Sum
snapchat_sessions	Total number of Snapchat sessions	Count
snapchat_time	Total time spent on Snapchat sessions	Sum
linkedin_sessions	Total number of LinkedIn sessions	Count
linkedin_time	Total time spent on LinkedIn sessions	Sum
pinterest_sessions	Total number on Pinterest sessions	Count
pinterest_time	Total time spent on Pinterest sessions	Sum

3.2.2. Target variable

The distribution of the raw score is represented in the Figure 3. The participants' BDD score was calculated by taking the average of their replies to all of those 11 questions. To make the target variable binary the following rule was applied: if the mean value of the BDD score was 1 or 2, the value was set to 0, if the mean value was 3,4 or 5 it was set to 1. The Figure 4 represents the distribution of the target variable after binarizing it.

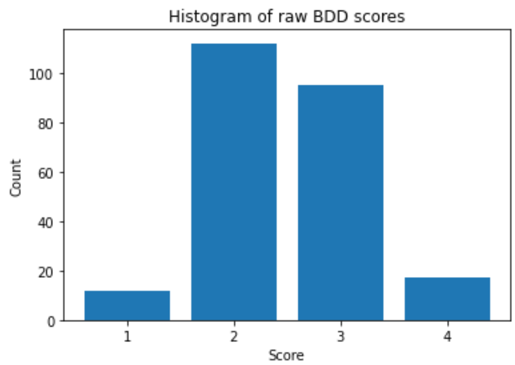


Figure 3. Distribution of raw scores BDD

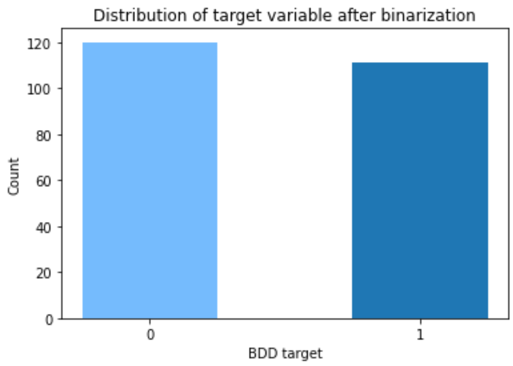


Figure 4. Distribution of target variable

The reason behind converting the data from the Likert scale into a binary format is that a score of 2.5 on a BDD level is not relevant. Instead, it should note if a certain person has BDD or not, as a yes (1) or no (0) answer. Converting the data in into a binary class supports that process.

3.3. Exploratory Data Analysis

3.3.1. Analysis of the independent variable

The mean score for BDD was 2.495 (SD = 0.705) before converting it to a binary class. The correlations between a binary variable and continuous variable were computed using Point biserial correlation coefficient. It was found that dependent variables are not correlated with target variable (See Appendix A).

3.3.2. Analysis of the dependent variables

Table 2 represent the 5 most popular categories among students in the period of February and June 2020. The applications as communication (40.68%) and Social (18.58%) are both higher than tools applications (7.76%).

Table 2. Application categories

Category	Count of sessions	Percentage
Communication	3104262	40.68%
Social	1417788	18.58%
Tools	591970	7.76%
Productivity	572862	7.51%
Photography	270343	3.54%

Figure 3 represents the frequency of social media application usage by session count. According to the figure 3 Facebook is the most popular social media platform among Tilburg University students, followed by Instagram, Snapchat, and LinkedIn. Based on the illustrated data the additional variables were formed as a result of the following applications. These variables were used to count the number and total time spent of each application's sessions.

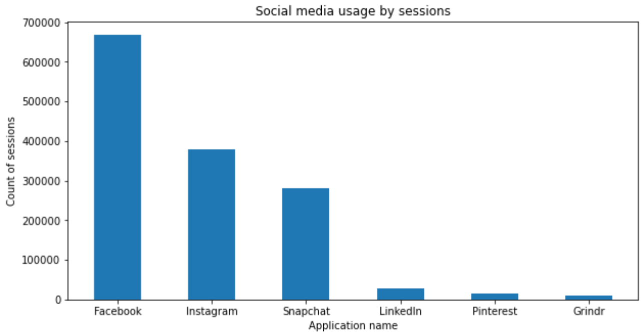


Figure 5. Social media usage by sessions

Table 3 presents the descriptive statistics for social media usage. The table shows that average time for Instagram time usage is relatively higher than Snapchat, LinkedIn, Pinterest or even Facebook. When the frequency of session is considered, the mean is higher for Snapchat session than Facebook session, and the time spent on Facebook is higher than Snapchat.

Table 1. Descriptive statistics of features

	MEAN	SD	SKEWNESS
insta_time	4254.093	4787.884	1.499
insta_sessions	2268.203	2639.617	1.667
fb_time	1361.496	2837.780	3.603
fb_sessions	674.922	1336.498	3.414
snap_time	1056.329	1816.436	3.581
snap_sessions	1337.839	2263.193	2.755
linked_sessions	8.761	49.057	7.705
linked_time	9.638	54.269	8.109
pinterest_time	185.829	1158.934	10.583
pinterest_sessions	72.502	456.198	10.712

3.4. Feature importance

The Random Forest classifier is one of the most used techniques due its simple and effective way of evaluating the importance of the feature in a dataset and it will be used in this research to determine which features attribute the most to a high performance of a model.

3.5. Feature selection techniques

When developing a machine learning models only most important features should be selected. The model performance might decrease if irrelevant features are used. Feature selection eliminates unnecessary features that do not have any effect on the model performance. Table 4 represents the number and the features chosen by given feature selection technique.

3.4.1. Full set

The full set contains all the features that were collected from the application data.

3.4.2. Theory-driven

10 Theory-driven features are chosen based on a theoretical knowledge based on social media influence on mental health disorders explained in the Literature Review.

3.4.3. ANOVA f-test

F-test was utilized in this research, giving an f-score by calculating the ratio of variances. The features were ranked based on their f-score values (Dhanya, Paul, Akula, Sivakumar, & Nair, 2020). ANOVA is particularly useful when one variable is numerical and the other is categorical, as in numerical input data and a classification goal variable in a classification problem.

3.4.4. Mutual information

Mutual information is the application of information gain, usually utilized in the building of decision trees for feature selection. Mutual information was measured between two features and quantifies the loss in uncertainty for one variable when the other variable's value is known (Vergara & Estevez, 2014). Mutual information as ANOVA f-test is mostly used when input and output data is categorical. It may, however, be used with numerical input and categorical output.

Table 4. Selected features using different selection techniques

Feature Selection	**Total Features**	**Selected Features**
Full set	18	total_sessions, time_sessions, sns_mean_time, sns_max_time, sns_min_time, insta_time, insta_sessions, fb_time, fb_sessions, snap_time, snap_sessions, sns_sessions, sns_time, linkedin_sessions, linked_time, pinterest_sessions, pinterest_time
ANOVA	15	total_sessions, time_sessions, sns_mean_time, sns_max_time, insta_time, insta_sessions, fb_time, fb_sessions, snap_time, snap_sessions, sns_sessions, sns_time, linkedin_sessions, linkedin_time, pinterest_sessions
Mutual Information	9	total_sessions, sns_mean_time, insta_time, insta_sessions, fb_sessions, sns_sessions, sns_time, pinterest_time, pinterest_sessions
Theory-driven	10	insta_sessions, insta_time, fb_time, fb_sessions, snap_time, snap_sessions, linkedin_sessions, linkedin_time, pinterest_sessions, pinterest_time

3.5. SMOTE

The target class in this report was almost equally represented (See Figure 4). However, slightly imbalanced data might provide a difficulty for models to predict. The issue with training the model on an unbalanced dataset is that the model will be biased solely towards the dominant class. In order to avoid this problem, Synthetic Minority Over-Sampling Technique (SMOTE) algorithm will be utilized. By randomly creating additional minority sample points to enhance the imbalance rate to a given level, the SMOTE method can improve the classification effect of unbalanced data (Wang, Dai, & Shen, 2021).

3.6. Cross-validation

Most previous research predicting mental health employed k-fold cross-validation to validate the models. As a result, prediction models with 10-fold cross validation are used in this research. To assess the performance of the models, the data is divided into two subsets: 70% training data and 30% testing data.

3.7. Algorithms

The supervised learning models as Logistic Regression, Random Forest, K-Nearest Neighbor and Support Vector Machines will be applied in this research. These machine learning algorithms seem to be well suited for binary classification and have been used in several studies related to mental health predictions (See Section 2.4). However, each algorithm has its own strength and weakness, the most important differences are outlined in the section below.

3.6.1. Logistic Regression

Logistic Regression is the simplest model to implement of all the models discussed above since it does not require tuning many hyperparameters. It is also more computationally efficient than SVM and Random Forest. Logistic Regression may investigate the relationship between several independent variables and a binary dependent variable, making it helpful for discovering characteristics associated with a certain result, such as emergence of mental health disorder.

3.6.2. Random Forest

Random Forest is an ensemble method that combines multiple decision trees to make predictions. The number of trees grown, as well as the complexity of each tree can make the training process computationally expensive when compared to other models, such as Logistic Regression, but less so when compared to SVMs. However, when the size of the forest is small, Random Forest is able to make fast predictions, because it does not require a lot of optimization. Another reason of choosing Random Forest is that it is less prone to overfitting than Logistic Regression, when working with small datasets, which is the case in this research.

3.6.3. K-Nearest Neighbor (KNN)

KNN is a simple and easy algorithm to implement, but its efficiency is strongly dependent on the choice of the number of nearest neighbors and the distance metric utilized.

KNN, like SVM, is a non-parametric algorithm. KNN makes predictions based on the majority class of the k-closest example, whereas SVM uses a decision boundary to maximize the margin between two classes. Also, KNN can deal with missing data since it depends on the similarity of the input data in the training set and doesn't require complete data for all input variables. This can be beneficial when considering the context of body dysmorphic disorder since many elements that contribute to the formation of this disorder are unknown.

3.6.4. Support Vector Machines (SVMs)

SVMs are strong models capable of handling non-linear decision boundaries, are resistant to overfitting and can handle high-dimensional data. These characteristics make SVM suitable for predicting mental health disorders as mental health can be influenced by a wide range of factors and these factors can interact in complex ways. On the other hand, SVM are generally considered as the most computationally expensive model. SVM needs to solve an optimization problem to determine a decision boundary, which typically takes a long time, especially when the number of features is large or when data is not linearly separable. Due to complexity of decision boundaries used in SVMs, the final model, like Random Forests, might be difficult to interpret.

3.6.5. Alternatives

The Naïve Bayes is one of the mostly used models in the binary classification. However, it won't be used in these studies since it did not perform very well on predicting different mental health disorders compared to other models mentioned in the Literature Review (see section 2.4).

3.6. Standardization

The scaling of features is an important stage in modeling algorithms using datasets. The data that is utilized for modeling is obtained through questionnaires and phone logs. As a result the acquired data contains features of different dimensions and sizes. The various sizes of data features could have a negative impact on dataset modeling. One of the results might be a skewed prediction results in terms of misclassification error and accuracy score.

Standardization is scaling approach that converts the statistical distribution to make it scale free using sklearn in Python. After the features are scaled, the models are built sequentially.

3.7. Normalization

The normalization for example, on features such as week of the days, was implemented since the data had different scales and one model used in this research, K-Nearest Neighbors, does not make any assumptions about the distribution of the data.

3.8. Hyperparameter tunning

The hyperparameter tuning is completed before evaluation of the model. The grid search is the most fundamental hyperparameter tuning. It simply generates a model for each potential combination of all hyperparameters that are provided, then evaluate each model and choose the parameters that deliver the best results for the particular model. The hyperparameters that were tuned for the models are listed in the Table 5.

Table 5. Hyperparameters tuning

Model	**Hyperparameter**
KNN	*leaf size:* the minimum number of points in a given node, *p:* optimal distance, *n_neighbors:* optimal number of neighbors
Support Vector Machines	*c:* regularization parameter
Random Forest	*n_estimators:* number of trees, *max_features:* the number of features to consider for the best split, *max_depth:* maximal depth of a tree, *min_samples_leaf:* minimal number of samples required to be at a leaf node, *bootstrap:* whether used or not
Logistic Regression	*penalty:* penalty term*, c:* inverse of regularization strength, *solver:* optimalization problem

3.9. Evaluation metrics

The first metric to evaluate the model’s performance was accuracy. Accuracy is perceived as one of the most used evaluation metrics in classification models. The proportion of properly categorized instances in the total set of examples is defined as accuracy. Furthermore, Area under the curve (AUC) was the second metric to evaluate the model’s performance. The AUC is undoubtedly one of the most efficient ways to assess how well a binary classification model works.

3.9.1. Baseline

To evaluate the model performance the baseline is essential. Due to lack of research on applying machine learning models to predict body dysmorphic disorder, the dummy classifier was used to establish the baseline for each of 4 feature sets.

3.10 Software

Python is used for coding, data analysis, and data visualization. Table 6 lists the packages utilized in this research, along with their source and version number.

Table 2. Libraries used in the research

Package name	Version	Source
Pandas	1.3.4	(The pandas development team, 2020)
NumPy	1.20.3	(Harris, Millman, & van der Walt, 2020)
scikit-learn	1.1.1	(Pedregosa, Varoquaux, Gramfort, Michel, & Thirion, 2011)
Matplotlib	3.4.3	(Hunter, Dale, Firing, & Droettboom, 2012)
Seaborn	0.11.2	(Waskom, 2021)

4. Results

This section presents the research's finding in four parts. First feature importance, followed by feature selection techniques, classification model performance and finally an examination of the models' performance when split by gender subset.

4.1. Feature Importance

This section highlights the importance of each feature in order to improve the model performance. The features are represented on the x-axis in Figure 6 and are arranged in descending order with the most important feature appearing first. To calculate the feature importance score (y-axis), the Random Forest technique was used. From the results it can be concluded that the most important feature when predicting BDD is number of total sessions run on the phone, and the least important feature is social networking service minimal time.

When looking at social media applications, the most important social media applications when predicting BDD is Instagram, Facebook, and Snapchat. The least important social media applications are LinkedIn and Pinterest.

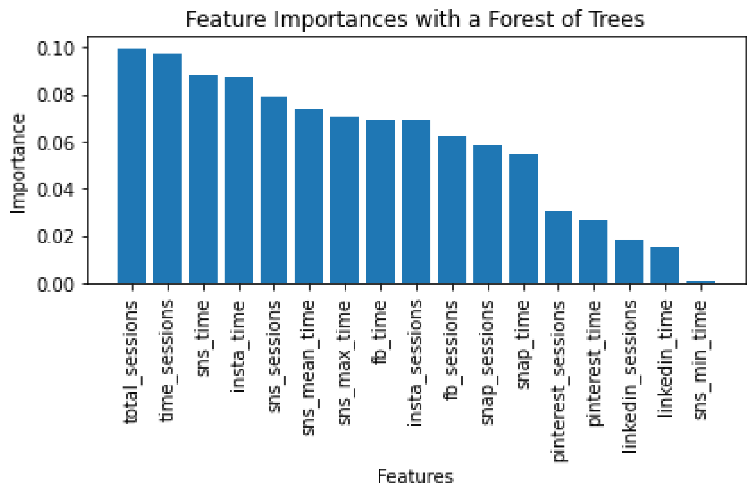


Figure 6. Feature Importance Random Forest

4.2. Feature Selection Techniques

In this section the results of the feature selection technique accuracy will be presented. Table 7 represents the accuracy results from the chosen feature selection methods such as Full set, ANOVA, Mutual Information and Theory Driven features. These methods were tested on fours chosen models. Since, theory-driven approach might result in the omission of important features that are not recognized in the theoretical framework, using other data-driven feature selection methods in addition to theory-driven feature selection was considered. The best performing method was Theory Driven, where it outperformed the other methods on all models except for Linear Regression. The worst performing feature selection method is Mutual Information (MI), scoring the lowest on KNN and LR. From these results it can be concluded that the best feature subset for predicting BDD is the Theory Driven method. As a result, the theory driven technique with 10 features would be used as the feature selection process for this study.

Table 7. Accuracy of models with feature selection technique

	Feature selection			
Models	FULL	ANOVA	MI	THEORY-DRIVEN
KNN	0.528	0.528	0.442	0.542
RF	0.472	0.500	0.486	0.557
SVM	0.486	0.486	0.527	0.541
LR	0.500	0.500	0.472	0.458

4.3. Model Performance

In this section, classification performance for the Theory-driven subset of features described in section 3.5 will be presented. Table 8 illustrates the accuracy and the area under the curve of each of the classification models. The baseline was achieved by using dummy classifier as described in the Methodology section 3.9.1. The results show that Linear Regression (45.8%) has the lowest performance among all models and KNN (54.2%), Random Forest (55.7%) and SVM (54.1%) performed above the baseline. Surprisingly, Random Forest outperformed all other models and achieved the highest accuracy of 55.7%, it will be discussed further in more detail in the Discussion section.

Table 8. Model performance

	Evaluation metrics	
	Baseline 0.519	
Models	**Accuracy**	**AUC**
KNN	0.542	0.517
RF	0.557	0.468
SVM	0.541	0.472
LR	0.458	0.550

Figure 7 and Figure 8 show the error analysis for the two best performing models (KNN and RF) trained on social media data based on predicted labels and true labels. There are 2 possible outcomes, 0 meaning not having BDD and 1 having BDD. Both classifiers, KNN and Random Forest made a total of 79 predictions (e.g., 79 students were being tested for the presence of BDD). Out of these 79 cases, the KNN classifier predicted 38 cases correctly, 13 true positives, and 25 true negatives. Random Forest showed 39 correct predictions of which 23 true negatives and 16 true positives. In the dataset, 28 participants showed signs of BDD and 42 did not. Therefore it can be concluded that both models did not perform very accurately, but when comparing the 2 models, Random Forest showed the lowest error when predicting BDD being present and KNN showed the lowest error when predicting BDD not being there.

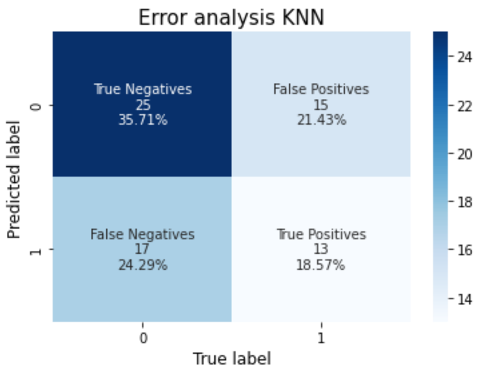


Figure 7. Error analysis KNN

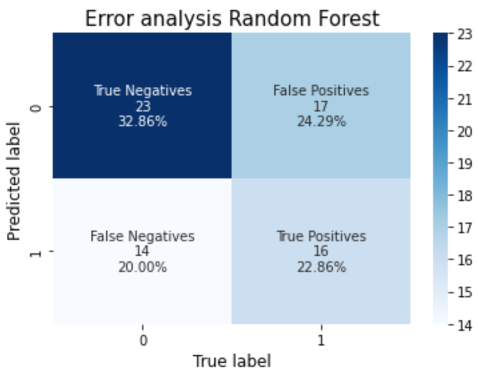


Figure 8. Error analysis Random Forest

4.4. Disparate groups analysis

The classification of each gender was individually analyzed to gain a better understanding of why the Random Forest model did not perform very well. Table 9 and Table 10 show the performance of the Random Forest on the test set separated in the subsets for men and women. When splitting the test set by gender, it is found that the model performs differently on men and women. The improvement over the baseline is much larger for women +20.8% than it is for men +17%.

As seen in the previous results, the feature selection technique driven by theory achieved the highest accuracy when predicting body dysmorphic disorder, however when testing the Random Forest model split by gender, the highest accuracy (82.4%) is achieved on female subset with features without any feature selection techniques.

Table 9. RF accuracy test set women

	FULL	ANOVA	MI	THEORY-DRIVEN
Baseline Female	0.528			
Female	0.824	0.800	0.784	0.736

Table 10. RF accuracy test set men

	FULL	ANOVA	MI	THEORY-DRIVEN
Baseline Male	0.575			
Male	0.745	0.764	0.726	0.745

There are several reasons for poor classification between men and women. One of them is that the model has learned to make predictions based on gender-specific patterns in the data, rather than general patterns that apply to both men and women. Table 11 represents the distribution of gender on training data in Random Forest. The model is trained on data where men are underrepresented, which may be less accurate when making predictions for men, even if it performs well on women.

Table 11. Training data distribution

Training data		
Gender	Men	Woman
Distribution	75	86

Another reason could be that the model is bias towards a certain group. After analysing the skewness of the subsets, it can be concluded that the training data is disproportionately skewed towards women resulting the model to predict women's characteristics more accurately (See Appendix) There are several techniques to mitigate the biases in the data such as debiasing methods, oversampling minorities, or underrepresented groups. They can help to improve the model's performance on these subgroups.

5. Discussion

This section of the research outlines the results and the answers to the research questions in detail and is divided into four separate sections. Each section discussing the research questions individually, and comparing it to the research done in the literature review. In the last section the research limitations are described, and possible further research proposals are made.

5.1. Features

The analysis of feature importance using the Random Forest selection technique revealed that the most important features when predicting Body Dysmorphic Disorder are the total number of sessions run on their phone, and the time spent during those sessions. When looking at different social media platforms, the most important features are time spent on Instagram, time spent on social networking services and time spent on Facebook.

When comparing the results to literature, … claims that certain social media platforms have a greater impact on appearance related disorders than others. The author claims that especially social media platforms that focus self-presentation and confirmation seeking activities have a large negative effect on someone's self-appearance. Moreover, due to the high diversity of pictures available on those self-presentation platforms, a lot of users said it created a feeling of constant comparison and competitiveness with peers.

This is in line with another recent study which reported that photo-based platforms, where users are able to upload pictures, are considered to have a negative effect on the peoples' appearance since they place a greater emphasis on physical beauty. The dilemma is compounded by the omnipresent use of filters that have recently become a must-have for every social media post. As the results and the literature conclude, these seemingly harmless pictures on self-presentation platforms and the use of filters, which allow people to improve their looks, have several negative effects on the well-being and mental health disorders such as BDD. These finding are not surprising since when exploring the research data it was found that Facebook, Instagram and Snapchat were the most used social media applications among Tilburg University Students (see section 3.3).

5.2. Feature selection method

The results when testing four different methods (Full set, ANOVA, Mutual Information and Theory Driven) on four chosen models showed that the theory driven method outperformed the other methods on all models except for Logistic Regression. As a result, the theory driven technique with 10 features performs best to predict BDD and led to a higher accuracy score of the model.

The well performing Theory Driven method was surprising and contradicted the literature. Zulfiker et al. (2021) researched predicting depression based on the Burns Depression Checklist (BDC) data by using selection techniques such as the ANOVA f-test, minimum redundancy and maximum relevance (mRMR), and the Boruta feature selection algorithm. It was discovered that the ANOVA feature selection method surpassed the other techniques and achieved the highest classification accuracy of 92.56% with the Adaboost classifier.

Moreover, Ernala et al. (2019) study about mental health was focused on using several social media data, containing linguistic information from Facebook and Twitter accounts from the participants activity such as comments and shared posts. The researcher found that from the used methods, the highest performing methods was ANOVA that achieved 75% accuracy. This is in line with the study from Zulfiker et al. (2021), but also contradicts this research.

A reason for this difference in results could be that ANOVA features selection methods are typically most effective when the features are continuous, the classes are well separated, and the data is normally distributed. In Zulfiker's et al. (2021) research, all the variables were transformed to continuous variables and SMOTE was applied to balance the data, which is something that was also done in this research. However, it did not lead to the same results.

To conclude, the theory-driven approach in this research resulted in high accuracy. However, when applying it in different research with different dataset it may result in the omission of important features since the theoretical framework was mainly focused on social media influence on BDD. So, it is necessary to consider using other data-driven feature selection methods in addition to proposed theory-driven features in this study.

5.3. Models

Despite the small amount of dataset related to social media usage derived from participant's phone data collection it was possible to identify body dysmorphic disorder state (i.e. having BDD or not) with good predictive performance using machine learning models. Interestingly, Random Forest, a tree-based classifier outperformed linear classifiers such as Support Vector Machines and Logistic Regression, conforming the presence of nonlinear correlation between social media usage and body dysmorphic disorder. Random Forest achieved the highest accuracy (55.7%) among the four models tested in the research to make predictions.

When comparing these results to the literature, the research of Angskun et al.(2022) and Almouzini et al. (2019) both showed that the most accurate model when predicting mental health disorders is Random Forest. However, Angskun et al.(2022) saw an accuracy of 74.35% and Almouzini et al. (2019) saw an accuracy of 73.5% which is far higher than the accuracy of the Random Forest model in this research (55.7%). The differences between their research and this research are that Angskun et al. (2022) build a depression detection model, training it with demographic factors and tweets instead of time spent on, and the number of sessions run on social media applications. But we both used the score of a self-reported Health Questionnaire for the outcome variable. Almouzini et al. (2019) used a totally different approach where they used Arabic tweets to build a model classifier that can detect whether a tweet represents any signs of depression or not.

It can be concluded that Random Forest performed the best on small data set, in Angskun 230 participant and in Almouzini 89 participants. However, the differences between the accuracy of the same model could be that both researchers used a linguistic approach, where this research uses continuous variables. Additionally, the features used in this research might be irrelevant to the problem and result in overfitting, which is when the model performs well on training data but badly on test data. By employing more relevant features, the other research may have prevented overfitting.

A surprising result is that the Support Vector Machines (SVMs) did not perform well in this research (54.1%) in comparison to the literature. Bakar et al. (2021) tried to detect depression level from tweet data from more than 10,314 random users by focusing on a more linguistic approach. This resulted in an accuracy of 78.27% on the Support Vector Machine. Another research of Gkotsis et al. (2017) analyzed Reddit posts from 32,280 users and classifiers were built to detect and categorize posts concerning mental illness based on 11 disease themes such as depression, self-harming etc. SVM saw an accuracy of 85.87%, making it the highest performing model.

The differences in accuracy between this research and the literature might be because SVM is particularly well-suited for problems with a high-dimensional features, such as text data. SVMs have been found to perform better in previous research based on text data, because they can effectively handle high-dimensional feature spaces and can find the most important features to separate the different classes. Since this research only uses continuous variables and no text data, the accuracy of the SVM is lower.

5.4. Gender

The disparate groups analysis revealed that when splitting the test set by gender, the model performs differently on men and women. The improvement over the baseline is much larger for women, +20.8% than it is for men, +17%. When comparing this to literature, Henn et al. (2019) suggests that in general, more women suffer from a negative body image than men. (Pikoos, Rossell, Tzimas, & Buzwell, 2021) adds that promoting an unrealistic idea of beauty may also lead to the development of BDD in women. The research suggests that there is a chance that more women than men suffer from a negative body image which can result in BDD. Another study of .. claims that hormonal influence also can play a role on women suffering quicker from BDD than men. The researchers claim that hormonal changes during puberty or menopause can change the way a woman looks at her body, which may increase the chance of BDD.

When looking at more data science related literature, the data size of training and testing set can affect the model's performance. When the training set is too small, for example when splitting the data into gender, the model might not be able to understand patterns and / or hyperparameters in the data. When the test set is too small, accuracy of the model could be inaccurate. This could be the case in this research, since the dataset used is already relatively small when comparing it to other mental health disorder studies.

Another reason could be bias in the data. A source for the higher improvement over the baseline for women could be that the men in the dataset were unrepresented. More women than men filled in the questionnaire, which means that there is a chance that a model recognizes patterns in the data of the women quicker than in the subset for men.

5.5. Further research and limitations

The developed model has several limitations. This study created a model using information from participants who gave consent to disclosure in exchange for study credits. Furthermore, the information employed in this study is limited to a certain degree, such as time spent on certain applications and self-reported mental health well-being score. The body dysmorphic disorder detection needs more information than provided by the log application. These factors may have impact on how BDD is analyzed.

For future research, some improvements in the data might be required. To begin, the number of participants in this study contained only 231 students from Tilburg University. The number of students for the next research should be higher to improve the accuracy of the models. Furthermore, collecting data from different universities might help to reduce the bias.

Secondly, to create a better model, other variables such as the activity of a certain application or information regarding that application like number of followers and likes might need to be explored to increase the accuracy of the model. Moreover, the model's development process could be improved by integrating additional data such as pictures or linguistic features into the current dataset. Finally, to keep the research relevant, it might be a good idea to add new data from social networks such as TikTok and BeReal into the dataset, as they are growing to be the biggest social media platforms.

6. Conclusion

One of the most neglected mental health disorders by researchers that is prominent among students is Body Dysmorphic Disorder (BDD). That is why the goal of this research is to get a deeper understanding of BDD using Machine Learning models, in relation to extensive social media application usage to contribute to decreasing the number of students with BDD.

This research contributes to the research of Body Dysmorphic Disorder by discovering to what extent Machine Learning Techniques can help identify the sources of

BDD. As discussed in the results, the Random Forest model has the highest accuracy mean of 55.7%. However, this research also showed that making predictions using Machine Learning models can be complex, and is very dependent on the chosen dataset and features selected.

Therefore, this study provides a good baseline for future studies to use when trying to predict Body Dysmorphic Disorders using machine learning algorithms based on social media data.

Acknowledgements

I would like to thank my supervisor Mr. Hendrickson for providing me with the dataset that I was able to use to write this research, and my family and friends for support during these times.

References

Bickham, D. S., Hswen, Y., & Rich, M. (2015). Media use and depression: exposure, household rules, and symptoms among young adolescents in the USA. *International Journal of Public Health*, 147-155 https://doi.org/10.1007/s00038-014-0647-6.

Lin, Y.-H., Pan, Y.-C., Lin, S.-H., & Chen, S.-H. (2017). Development of short-form and screening cutoff point of the Smartphone Addiction Invetory (SPAI-SF). *Methods Psychiatr Res*, 1-6 https://doi.org/10.1002/mpr.1525.

Choi, J., Rho, M., Kim, Y., Yook, I., & Yu, H. (2017). Smartphone dependence classification using tensor factorization. *PloS ONE*, https://doi.org/10.1371/journal.pone.0177629.

Montag, C., Blaszkiewicz, K., Lachmann, B., Andone, I., Sariyska, R., Trendafilov, B., . . . Markowetz, A. (2014). Correlating personality and actual phone usage: Evidence from psychoinformatics. *Journal of Individual Differences*, 158-165 https://doi.org/10.1027/1614-0001/a000139.

Chancellor, S., Baumer, E. P., & De Choudhury, M. (2019). Who is the "Human" in Human-Centered Machine Learning. *The Case of Predicting Mental Health from Social Media*, 1-31 https://doi.org/10.1145/3359249.

Balakrishna, S., Gopi, Y., & Solanki, V. (2022). Comparative analysis on deep neural network models for detection of cyberbullying on Social Media. *Ingeniería Solidaria*, 1-33 https://doi.org/10.16925/2357-6014.2022.01.05.

Priya, A., Garg, S., & Tigga, N. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Computer Science,, 167*, 1258–1267 https://doi.org/10.1016/j.procs.2020.03.442.

Malhotra, A., & Jindal , R. (2022). Deep learning techniques for suicide and depression detection from online social media: A scoping review. *Applied Soft Computing, 130*, https://doi.org/10.1016/j.asoc.2022.109713.

Birnbaum, M., Ernala, S., Rozvi, A., De Choudhury, M., & Kane, J. (2017). A Collaborative Approach to Identifying Social Media Markers of Schizophrenia by Employing Machine Learning and Clinical Appraisals. *Journal of Medical Internet Research*, https://doi.org/10.2196/jmir.7956.

Ware, S., Yue, C., Morillo, R., Lu, J., Shang, C., Bi, J., . . . Wang, B. (2020). Predicting depressive symptoms using smartphone data. *Smart Health*, https://doi.org/10.1016/j.smhl.2019.100093.

Rashid, H., Mendu, S., Daniel, K., Beltzer, M., Teachman, B., Boukhechba, M., & Barnes, L. (2020). Predicting Subjective Measures of Social Anxiety from Sparsely Collected Mobile Sensor Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1-24 https://doi.org/10.1145/3411823.

Opoku, A. K., Terhorst, Y., Vega, J., Peltonen, E., Lagerspetz, E., & Ferreira, D. (2021). Predicting Depression From Smartphone Behavioral Markers Using Machine Learning Methods, Hyperparameter Optimization, and Feature Importance Analysis: Exploratory Study. *JMIR MHealth and UHealth*, https://doi.org/10.2196/26540.

Flygare, O., Enander, J., & Andersson, E. (2020). Predictors of remission from body dysmorphic disorder after internet-delivered cognitive behavior therapy: a machine learning approach. *BMC Psychiatry*, https://doi.org/10.1186/s12888-020-02655-4.

Rashid, H., Mendu, S., Daniel , K., Beltzer, M., Techman, B., Boukhechba, M., & Barnes, L. (2020). Predicting Subjective Measures of Social Anxiety from Sparsely Collected Mobile Sensor Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1-24 https://doi.org/10.1145/3411823.

Priya, A., Garg, S., & Tigga, N. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Computer Science*, 1258–1267 https://doi.org/10.1016/j.procs.2020.03.442.

Kumar, P., Garg, S., & Garg, A. (2020). Assessment of Anxiety, Depression and Stress using Machine Learning Models. *Procedia Computer Science*, 1989–1998 https://doi.org/10.1016/j.procs.2020.04.213.

Mohd Shafiee, N., & Mutalib, S. (2020). Prediction of Mental Health Problems among Higher Education Student Using Machine Learning. *International Journal of Education and Management Engineering*, 1-9 https://doi.org/10.5815/ijeme.2020.06.01.

Sekulic, I., & Strube, M. (2020). Adapting Deep Learning Methods for Mental Health Prediction on Social Media.

Benton, A., Mitchell, M., & Hovy, D. (2017). Multitask learning for mental health conditions with limited social media data. *15th conference of the European Chapter of the Association for Computational Linguistics*, 152-162 https://aclanthology.org/E17-1015.

Huang, X., Zhang, L., Liu, T., Chiu, D., Zhu, T., & Li, X. (2014). Deecting Suicidal Ideation in Chinese Microblogs with Psychological Lexicons. *Computation and Language*, 1-6 https://doi.org/10.48550/arXiv.1411.0778.

Chancellor, S., Mitra, T., & De Choudhury, M. (2017). Recovery amid pro-anorexia: analysis of recovery in social media. *Proceedings of the SIGCHI conference on human factors in computing systems*, 2111–2123 https://doi.org/10.1145/2858036.2858246.

Saha, K., Chan, L., De Barbaro, K., Abowd, G., & De Choudhury, M. (2017). Inferring Mood Instability on Social Media by Leveraging Ecological Momentary Assessments. *ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*.

Gkotsis, G., Oellrich, A., & Velupillai, S. (2017). Characterisation of mental helath conditions in social media using informed Deep Learning. *Sci Rep*, https://doi.org/10.1038/srep45141.

Braithwaite, S., Giraud-Carrier, C., West, J., Barnes, M., & Hanson, C. (2016). Validation Machine Learning Algorithms for Twitter Data Against Established Measures of Suicidality. *Mental Health*, https://doi.org/10.2196/mental.4822.

Seabrook, E., Kern, M., Fulcher, B., & Rickard, N. (2018). Predicting Depression From Language-Based Emotion Dynamics: Longitudinal Analysis of Facebook and Twitter

Status Updates. *Journal of Medical Internet Research*, https://doi.org/10.2196/jmir.9267.

Richard, B., Marsch, L., Crosier, B., & Hassanpour, S. (2018). Exploring the Utility of Community-Generated Social Media Content for Detecting Depression: An Analytical Study on Instagram. *Journal of Medical Internet Research*, https://doi.org/10.2196/11817.

De Choudhury, M. (2015). Anorexia on Tumblr: A Characterization Study on Anorexia. *Proceedings of the 5th International Conference on Digital Health 2015*, https://doi.org/10.1145/2750511.2750515 .

Cohan, A., Young, S., Yates, A., & Goharian, N. (2017). Triaging content severity in online mental health forums. *Journal of the Association for Information Science and Technology*, 2675–2689 https://doi.org/10.1002/asi.23865.

Wang, T., Brede, M., Ianni, A., & Mentzakis, E. (2017). Detecting and Characterizing Eating-Disorder Communities on Social Media. In Proc. *10th ACM International Conference on Web Search and Data Mining (WSDM)*, 91-100 https://doi.org/10.1145/3018661.3018706.

Ive, J., Gkotsis, G., Dutta, R., Stewart, R., & Velupillai, S. (2018). Hierarchical neural model with attention mechanisms for the classification of social media text related to mental health. . *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic.*, https://doi.org/10.18653/v1/w18-0607.

Bjornsson, A., Didie, E., & Philips, K. (2010). Body dysmorphic disorder. *Dialogues in Clinical Neuroscience*, 221-232 https://doi.org/10.31887/dcns.2010.12.2/abjornsson.

Schneider, S., Turner, C., Storch, E., & Hudson, J. (2019). Body dysmorphic disorder symptoms and quality of life: The role of clinical and demographic variables. . *Journal of Obsessive-Compulsive and Related Disorders*, 1-5 https://doi.org/10.1016/j.jocrd.2018.11.002.

Hartmann, A., Cordes, M., Hirschfeld, G., & Vocks, S. (2019). Affect and worry during a checking episode: A comparison of individuals with symptoms of obsessive-compulsive disorder, anorexia nervosa, bulimia nervosa, body dysmorphic disorder, illness anxiety disorder, and panic disorder. *Psychiatry research*, 349–358 https://doi.org/10.1016/j.psychres.2018.12.132.

Krebs, G., Fernandez de la Cruz, L., & Mataix-Cols, D. (2017). Recent advances in understanding and managing body dysmorphic disorder. *Evidence-based mental health*, 71-75 https://doi.org/10.1136/eb-2017-102702.

Shoukat, S. (2019). Cell phone addiction and psychological and physiological health in adolescents. *EXCLI journal*, 47-50 .

Twenge, J., Martin, G., & Campbell, W. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of smartphone technology. 765–780 https://doi.org/10.1037/emo0000403.

Sarker, I. (2019). A machine learning based robust prediction model for real-life mobile phone data. *Internet of Things*, 180-193 https://doi.org/10.1016/j.iot.2019.01.007.

Sigurdarson, T. (2022). Social media usage by sessions. 1-24.

Bjornsson, A., Didie, E., & Philips, K. (2022). Body dysmorphic disorder Trastorno dismórfico corporal Trouble dysmorphophobique. *Dialogues in Clinical Neuroscience*, 221-232 https://doi.org/10.31887/DCNS.2010.12.2/abjornsson.

Veale, D. (2004). Body dysmorphic disorder. *Postgraduate Medical Journal*, 67-71.

Dhanya, R., Paul, I., Akula, S., Sivakumar, M., & Nair, J. (2020). F-test feature selection in Stacking ensemble model for breast cancer prediction. *Procedia Computer Science*, 1561-1570 https://doi.org/10.1016/j.procs.2020.04.167.

Vergara, J., & Estevez, P. (2014). A review of feature selection methods based on mutual information. *Neural Comput & Application*, 175-186 https://doi.org/10.1007/s00521-013-1368-0.

Wang, S., Dai, Y., & Shen, J. (2021). Research on expansion and classification of imbalanced data based on SMOTE algorithm. *Sci Rep*, https://doi.org/10.1038/s41598-021-03430-5.

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, 12–18 https://doi.org/10.11613/bm.2014.003.

Schnebly, J., & Sengupta, S. (2019). Random Forest Twitter Bot Classifier. *Computing and Communication Workshop and Conference*, https://doi.org/10.1109/ccwc.2019.8666593.

Chomboon, K., Chujai, P., Teerarassammee, P., Kerdprasop, K., & Kerdprasop, N. (2015). An Empirical Study of Distance Metrics for k-Nearest Neighbor Algorithm. *The Proceedings of the 2nd International Conference on Industrial Application Engineering 2015*, https://doi.org/10.12792/iciae2015.051.

Huang, S., Cai, N., Pacheco, P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer genomics & proteomics*, 41-51 https://doi.org/10.21873/cgp.20063.

Signh, A., & Veale, D. (2019). Understanding and treating body dysmorphic disorder. *Indian journal of psychiatry*, 131-135 https://doi.org/10.4103/psychiatry.IndianJPsychiatry_528_18.

Henn, A., Taube, C., Vocks, S., & Hartmann, A. (2019). Body Image as Well as Eating Disorder and Body Dysmorphic Disorder Symptoms in Heterosexual, Homosexual, and Bisexual Women. *Frontiers in Psychiatry*, https://doi.org/10.3389/fpsyt.2019.00531.

The pandas development team. (2020). pandas-dev/pandas: Pandas. https://doi.org/10.5281/zenodo.3509134.

Harris, C., Millman, K., & van der Walt, S. (2020). Array programming with NumPy. *Nature*, 357-362 https://doi.org/10.1038/s41586-020-2649-2.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., & Thirion, B. (2011). Scikit-learn: Machine Learning in python. *Journal of Machine Learning Research*, 2825-2830.

Hunter, J., Dale, D., Firing, E., & Droettboom, M. (2012). The Matplotlib development team.

Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 3021 https://doi.org/10.21105/joss.03021.

Eichstaedt, J., Smith, R., & Merchant, R. (2018). Facebook language predicts depression in medical records . *Psychological and cognitive sciences*, 11203-11208 https://doi.org/10.1073/pnas.1802331115.

Chung, H., Park, C., Kang, W., & Lee, J. (2021). Gender Bias in Artificial Intelligence: Severity Prediction at an Early Stage of COVID-19. *Frontiers in Physiology*, https://doi.org/10.3389/fphys.2021.778720.

Cuzzolaro, M., & Nizzoli, U. (2018). Enrico Morselli and the Invention of Dysmorphophobia . *Body Image, Eating and Weight*, 85-95 https://doi.org/10.1007/978-3-319-90817-5_6.

Zulfiker, M., Kabir, N., Biswas, A., Nazneen, T., & Uddin, M. (2021). An in-depth analysis of machine learning approaches to predict depression . *Current Research in Behavioral Sciences* , https://doi.org/10.1016/j.crbeha.2021.100044.

Schieber, K., Kollei, I., de Zwaan, M., & Martin , A. (2015). Classification of body dysmorphic disorder — What is the advantage of the new DSM-5 criteria? *Journal of Psychosomatic Research*, 223-227 https://doi.org/10.1016/j.jpsychores.2015.01.002.

Veale, D., Gledhill, L., Christodoulou, P., & Hodsoll, J. (2016). Body dysmorphic disorder in different settings: A systematic review and estimated weighted prevalence. *Body Image*, 168-186 https://doi.org/10.1016/j.bodyim.2016.07.003.

Pikoos, T., Rossell, S., Tzimas, N., & Buzwell, S. (2021). Is the needle as risky as the knife? The prevalence and risks of body dysmorphic disorder in women undertaking minor cosmetic procedures. *Australian & New Zealand Journal of Psychiatry*, 1191-1201 https://doi.org/10.1177/0004867421998753.

Harrison, A., Fernández de la Cruz, L., Enander, J., Radua, J., & Mataix-Cols, D. (2016). Cognitive-behavioral therapy for body dysmorphic disorder: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review*, 43-51 https://doi.org/10.1016/j.cpr.2016.05.007.

Thakur, S., & Roy, R. (2020). Predicting mental health using smart-phone usage and sensor data. *Journal of Ambient Intelligence and Humanized Computing*, 9145–9161 https://doi.org/10.1007/s12652-020-02616-5.

Ernala, S., Birnbaum, M., Candan, K., Rizvi, A., Sterling, W., Kane, J., & De Choudhury, M. (2019). Methodological Gaps in Predicting Mental Health States from Social Media. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, https://doi.org/10.1145/3290605.3300364.

Almouzini, S., khemakhem, M., & Alageel, A. (2019). Detecting Arabic Depressed Users from Twitter Data. *Procedia Computer Science*, 257–265 https://doi.org/10.1016/j.procs.2019.12.107.

Angskun, J., Tipprasert, S., & Angskun, T. (2022). Big data analytics on social networks for real-time depression detection. *J Big Data*, https://doi.org/10.1186/s40537-022-00622-2.

Angskun, J., Tipprasert, S., & Angskun, T. (2022). Big data analytics on social networks for real-time depression detection. *Journal of Big Data*, 1-15 https://doi.org/10.1186/s40537-022-00622-2.

Bakar, F., Nawi, N., & Hezam, A. (2021). Predicting Depression Using Social Media Posts. *Journal of Soft Computing and Data Mining*, 39-48 https://doi.org/10.30880/jscdm.2021.02.02.004.

Sharma, M., John, N., & Sahu, M. (2020). Influence of social media on mental health. *Current Opinion in Psychiatry, Publish Ahead of Print*, https://doi.org/10.1097/yco.0000000000000631.

Fardouly, J., & Vartanian, L. (2016). Social Media and Body Image Concerns: Current Research and Future Directions. *Current Research and Future Directions. Current Opinion in Psychology*, 1-5 https://doi.org/10.1016/j.copsyc.2015.09.005.

Carey, R., Donaghue, N., & Broderick, P. (2014). Body image concern among Australian adolescent girls: The role of body comparisons with models and peers. *Body Image*, 81-84 https://doi.org/10.1016/j.bodyim.2013.09.006.

Karsay, K., Trekels, J., Eggermont, S., & Vandenbosch, L. (2021). "I (Don't) Respect My Body": Investigating the Role of Mass Media Use and Self-Objectification on Adolescents' Positive Body Image in a Cross-National Study . *Mass Communication and Society* , 57-84 https://doi.org/10.1080/15205436.2020.1827432.

Appendices and Supplementary Materials

Appendix A: Exploratory Data Analysis

Table A 1.Correlation with target variable

Variable	Correlation	P-value
Total_Sessions	-0.014	0.828
Time_sessions	-0.063	0.337
Sns_mean_time	-0.097	0.138
Sns_max_time	-0.070	0.288
Sns_min_time	0.068	0.299
Insta_time	0.010	0.873
Insta_sessions	0.016	0.806
Fb_time	-0.030	0.642
Fb_sessions	-0.061	0.351
Snap_time	-0.017	0.789
Snap_sessions	0.007	0.911
Sns_Sessions	-0.006	0.917
Sns_time	-0.014	0.830
Linkedin_sessions	-0.040	0.540
Linkedin_time	-0.042	0.524
Pinterest_time	-0.019	0.765
Pinterest_sessions	-0.015	0.811

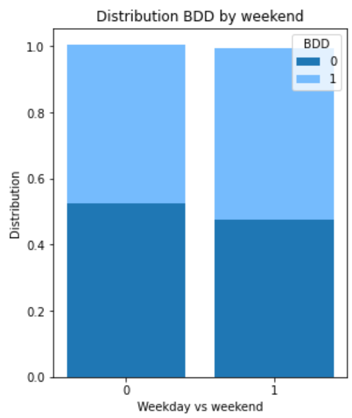
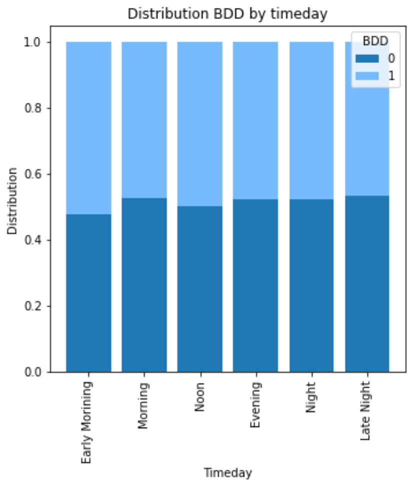


Figure A. 1. BDD on weekday vs weekend



Figures A. 2 BDD during the time day

Variable	Count sessions
Monday	148232
Tuesday	148337
Wednesday	151836
Thursday	144733
Friday	142517
Saturday	134280
Sunday	138336

Table A 2. Value count of SNS sessions grouped by the day of the week

Variable	Time	Count of sessions
Late-night	24:00 – 03:00	70894
Early Morning	04:00 – 07:00	15933
Morning	08:00 – 11:00	165959
Noon	12:00 – 15:00	244563
Evening	16:00 – 19:00	251686
Night	20:00 – 13:00	259236

Table A 3. Value count of SNS sessions grouped by daytime

Table 3. Optimal hyperparameters

Set number	KNN	Logistic Regression	Random Forest	SVM
1	leaf size: 1 p: 1 n_neighbors: 15	penalty: l1 c: 29.76 solver: liblinear	n_estimators: 10 max_features: sqrt max_depth: 4 min_samples_leaf: 1 min_samples_split: 5 bootstrap : False	c: 10000000000.0,
2	leaf size: 1 p: 1 n_neighbors: 12	penalty: l2 c: 0.001 solver: liblinear	n_estimators: 10 max_features: sqrt max_depth: 4 min_samples_leaf: 2 min_samples_split: 2 bootstrap : False	c: 100000000.0,
3	leaf size: 1 p: 2 n_neighbors: 2	penalty: l1 c: 29.76 solver: liblinear	n_estimators: 72 max_features: auto max_depth: 4 min_samples_leaf: 1 min_samples_split: 5 bootstrap : True	c: 10.0
4	leaf size: 1 p: 1 n_neighbors: 12	penalty: l2 c: 0.01 solver: liblinear	n_estimators: 64 max_features: auto max_depth: 4 min_samples_leaf: 1 min_samples_split: 5	c: 10

Training data Random Forest skewness on woman

total_sessions	0.630326
time_sessions	0.182683
sns_mean_time	1.085655
sns_max_time	1.552805
sns_min_time	0.000000
insta_time	1.234120
insta_sessions	1.458039
fb_time	2.645380
fb_sessions	2.843793
snap_time	4.049446
snap_sessions	2.698880
sns_sessions	1.578254
sns_time	0.995851
linkedin_sessions	7.585266
linkedin_time	7.028452
pinterest_time	6.979319

pinterest_sessions	7.027551

Training data Random Forest Skewness on man

total_sessions	1.028191
time_sessions	0.258116
sns_mean_time	2.383594
sns_max_time	0.962051
sns_min_time	5.996607
insta_time	2.118460
insta_sessions	1.426071
fb_time	5.322663
fb_sessions	4.473069
snap_time	2.722207
snap_sessions	2.426772
sns_sessions	1.305481
sns_time	1.248199
linkedin_sessions	6.003963
linkedin_time	6.171980
pinterest_time	8.528527
pinterest_sessions	8.000568