

Sl no	Paper Title and Author	Year of Publication	Aim and Objective	Uniqueness claimed	Outcome acheived	Algorithm/Tools/Platform used	Feature Extraction	Modality of data	Data Resource
1	Combining Sentiment Analysis Models Using Stacking Ensemble Learning Techniques on BIST30 Stocks by Mahmut Sami SIVRI	2024	Develop a stacking ensemble model to enhance sentiment analysis accuracy and robustness for BIST30 financial news.	Integrate diverse models (LSTM, BERT, Naive Bayes, SVM) in a stacking ensemble with a focus on BIST30 stocks, robust data strategies, and comparative analysis.	Enhance financial sentiment analysis with superior performance, generalization over baseline models with accuracy of 85%	Stacking ensemble with LSTM, BERT, Naive Bayes, SVM, and Logistic Regression, trained on 5,287 financial news articles (2020-2023) with an 80-10-10 split.	SVM uses TF-IDF with text preprocessing: cleaning, tokenization, stopwords removal, and lemmatization.	Financial news articles classified into Positive, Negative, and Neutral sentiment categories.	Thomson Reuters news articles
2	Predicting mental health using social media: A roadmap for future development by Ramin Safa, S. A. Edalatpanah, Ali Souroukhah	2022	A roadmap for analyzing and predicting mental disorders using social media data, covering collection, feature extraction, and prediction methods.	Comprehensive analysis of mental state assessment methods on social data, categorizing approaches and discussing challenges and future directions.	Conceptual framework for e-mental health research via social media, covering assessment strategies, data methods, algorithms, metrics, and future challenges.	Machine learning, NLP, and sentiment analysis using platforms like Twitter and Facebook, with tools like LIWC and frameworks such as CNN and LSTM.	Tools like LIWC, OpinionFinder, SentiStrength, VADER, Word2Vec, LDA, tf-idf, VGG-Net, and Imagga	Textual Data and Visual Content	myPersonality project CLPsych workshop data, eRisk workshops data, AutoDep dataset, SDCNL dataset, CAMS dataset
3	Detecting Depression and Mental Illness on Social Media: An Integrative Review by Sharath Chandra Guntuku, David B. Yaden, Margaret L. Kern, Lyle H. Ungar, Johannes C. Eichstaedt	2016	Reviewing recent studies and comparing approaches for predicting mental illness through social media analysis.	Comprehensive review comparing mental illness detection methods and prediction performances with clinical baselines.	Social media-based screening shows AUCs between 0.70 and 0.91, bridging clinician assessment and screening surveys with accuracy of 87%	Using Linear Regression, SVM, Neural Networks, and Random Forest with analysis tools like LIWC and LabMT.	LIWC and N-Gram	Textual Data	Survey and Social media
4	Single classifier vs. ensemble machine learning approaches for mental health prediction by Jetli Chung and Jason Teo	2023	Empirically evaluate and compare single classifiers and ensemble approaches for predicting mental health problems.	Comprehensive comparison of traditional algorithms, Deep Neural Networks, XGBoost, and ensemble methods for mental health prediction using survey data.	Gradient Boosting led with 88.80% accuracy, followed by Neural Networks (88.00%) and XGBoost (87.20%).	Logistic Regression, Gradient Boosting, Neural Networks, K-NN, SVM, Deep Neural Networks, XGBoost, and Ensemble Voting Classifier.	Extra Trees Classifier was used for feature selection to reduce overfitting	Survey	OSMI's 2014 Mental Health in Tech Survey measured workplace mental health attitudes in the tech industry, available under Creative Commons.
5	Survey of transformers and towards ensemble learning using transformers for natural language processing by Hongzhi Zhang and M. Omair Shafiq	2024	Compare and analyze transformer models across six NLP tasks, developing ensemble models to leverage their strengths.	Compare 5 transformer models across 6 NLP tasks, analyzing nuances and developing novel ensemble approaches.	Outperformed single classifiers with two ensemble models, identifying optimal combinations for specific NLP tasks with accuracy of 79%	BERT, XLNet, GPT2, RoBERTa, and ALBERT on Google Colab using Transformers, Deep Neural Networks, BERTopic, EncoderDecoderModel, and Seq2SeqTrainer.	Using transformer-based embeddings, word vectors, BERTopic for topic modeling, class-based TF-IDF, and contextual semantic representations.	Using transformer embeddings, BERTopic, TF-IDF, and contextual semantic representations for NLP.	Using datasets like Kaggle's coronavirus tweets, SQuAD1.1, Groningen Meaning Bank, CNN daily mail, disaster tweets, and Trump 2020 election speeches for various NLP tasks.
6	Ensemble of hybrid model based technique for early detecting of depression based on SVM and neural networks by Dip Kumar Saha, Tuhin Hossain, Mejd1 Safran, Sultan Alfahhood, M. F. Mridha & Dunren Che	2024	Develop an automated system to detect depression using psychological and sociodemographic characteristics.	Develop an automated system to detect depression using psychological and sociodemographic characteristics.	Developed an automated depression detection system with improved accuracy and reduced class imbalance.with accuracy of 99.39%	Algorithms like SVM, MLP, Hybrid DeprMVM, RFC, KNN, XGB with SMOTE, cluster sampling, and Scikit-learn for depression detection.	Using Label Encoder, Standard Scaler, and feature selection techniques for data preprocessing.	Survey data with 30 psychosocial/demographic predictors and 1 depression status target, including 25 BDC questions.	Dataset of 604 Bangladeshi participants (65.7% depressed, 34.3% not) collected from April to August 2020.
7	Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features by Hatoon AlSagri and Mourad Ykhlef (2023)	2023	Detect depression in Twitter users by analyzing tweet content and network behavior using machine learning techniques.	Incorporating tweet text, user behavior, and new features like activity categories, Dept_Sent lexicon, and synonyms for enhanced depression detection.	Developed a binary classification model with improved accuracy by combining user activities and tweets for better mental health detection with accuracy 82.5%	Using SVM, Naive Bayes, and Decision Tree algorithms with R 3.3, RStudio, Twitter API, WordNet, and R packages for data analysis.	Feature extraction using TF-IDF, WordNet, Information Gain, sentiment analysis tools, and R packages.	Text data (tweets, pronouns, sentiment words, depression terms) and user activity data (followers, posts, mentions, retweets, emojis, hashtags)	Collected over 300,000 tweets from 111 users via Twitter API, with manual verification of self-disclosed depression and data selection criteria.
8	Predicting Depression via Social Media" Authors & Year: Munmun De Choudhury, Michael Gamon, Scott Counts, Eric Horvitz	2013	Leveraging social media behavioral patterns to detect and predict Major Depressive Disorder (MDD) early.	Pioneering crowdsourced clinical data and multifaceted behavioral analysis for early depression prediction.	Achieved 70% accurate depression prediction, highlighting key behavioral and linguistic markers.	Utilized SVM with RBF kernel, PCA, and 10-fold cross-validation on Twitter data for depression analysis.	Employed LIWC, ANEW, and custom lexicons for linguistic, emotional, and medication-related feature extraction.	Analyzed Twitter posts, surveys (CES-D, BDI), demographics, and social network data for depression detection.	Collected data from 1,583 U.S.-based crowdworkers, analyzed 2M+ tweets and surveys (CES-D, BDI) with strict quality controls.
9	Predicting Mental Health Illness using Machine Learning Algorithms by Konda Vaishnavi	2022	Evaluate and compare machine learning techniques for accurate mental health issue prediction.	Unique comparison of five ML techniques for mental health prediction using diverse accuracy metrics and ROC curves.	Achieved 81.75% accuracy with stacking, with all classifiers showing strong ROC values (0.8–0.9).	Utilized Logistic Regression, K-NN, Decision Tree, Random Forest, and Stacking for mental health prediction.	Performed feature selection to reduce 27 attributes to 8 for mental health prediction.	Used a dataset with 27 attributes and 1259 entries, including text documents for mental health prediction.	The paper lacks details on dataset source, collection method, attributes, time period, and demographics.
10	Generalizability of Machine Learning to Categorize Various Mental Illness Using Social Media Activity Patterns by Ang, C.S. & Venkatachala, R.	2023	Explore linguistic patterns and cross-platform machine learning models for classifying mental health groups based on social media activity.	Improved accuracy by 2.11% with simpler cross-platform models for mental health classification on Twitter and Reddit.	Achieved 91.5% accuracy, with better Reddit model generalization and distinct linguistic patterns between platforms.	Used CNN, Word2Vec, NLTK, and Google Colab for Twitter and Reddit-based mental health classification.	Employed LIWC, Porter Stemmer, and NLTK for feature extraction in mental health classification.	Analyzed 606K Reddit posts and 23M tweets covering 6 mental health conditions from relevant subreddits and hashtags.	Used 23M Twitter tweets and 606K Reddit posts from mental health-related subreddits for cross-platform model training and testing.