Comments Mining With TF-IDF: The Inherent Bias and Its Removal.

K. SAI VARDHAN*
AIT-CSE (AIML), Chandigarh University, Mohali, India
saivardhankari@gmail.com
ORCID: 0009-0008-9558-8535

M. RAMA KRISHNA REDDY
AIT-CSE (AIML), Chandigarh University, Mohali, India
rk0010430@gmail.com
ORCID: 0009-0004-1747-0153

S. SAI VAIBHAV
AIT-CSE (AIML), Chandigarh University, Mohali, India
svaibhav6744@gmail.com
ORCID: 0009-0001-0060-1637

AASKARAN BISHNOI
AIT-CSE (AIML), Chandigarh University, Mohali, India
aaskaran.e15060@cumail.in
ORCID: 0009-0003-1853-2937

DAYAL CHANDRA *SATI*
AIT-CSE (AIML), Chandigarh University, Mohali, India
dayal.e13263@cumail.in
ORCID: 0000-0001-7736-8960

***Abstract*—The notable increase in user-generated content has made comment mining an essential component of natural language processing. The Term Frequency-Inverse Document Frequency (TF-IDF) is a key technique in this field that assigns weights to terms based on their occurrence in particular documents in relation to the overall corpus. Although it is advantageous, TF-IDF can lead to statistical bias, particularly in situations where language is closely intertwined, such as on social media, in political debates, product reviews, or educational evaluations. This bias stems from recurring themes in participant conversations or specialized terminology relevant to the domain, which can skew sentiment or topic analysis. Several research studies have tackled this challenge by suggesting improvements to TF-IDF to enhance sentiment analysis. For instance, TF-IDF has been utilized to measure polarity by analyzing the distinctiveness of phrases in political tweet assessments. Machine learning algorithms like SVM, when combined with TF-IDF, reached accuracies of up to 89.55% in evaluating IMDb reviews; further refinements, such as sentiment labeling using TextBlob, increased this accuracy to 92%. The fusion of TF-IDF with models like Naïve Bayes or Gradient Boosting enhanced effectiveness in mining reviews for products and employees while also emphasizing the necessity to tackle bias.**

Keywords*—TF-IDF, Comment Mining, Sentiment analysis, Bias correction, Machine learning, User-generated content, Polarity detection.

I. INTRODUCTION.

Social media, mostly network sites, has turned into a key way to talk for both people and groups. Moving from face-to-face to online talking makes a lot of text data. This is great for study in data jobs, business tech, and info systems. A common way to check this data is TF-IDF. It finds how key a word is by seeing how often it shows up in a text compared to how common it is in all texts. Yet, changes in how people write online can mess up TF-IDF results, mainly when you miss out on how people interact. This issue can mess up how we see comments and feelings, more so when we are setting up data for tests.

This work shows a new TF-IDF way to cut bias and better analysis, focusing on short texts like app blurbs and tags, and looks at other options like PCA when they fit. It digs into the limits of old TF-IDF in checking social media talks, where the setup of the talk is often missed. On platforms like Facebook, comments are seen as alone, not seeing how they link, mainly in threads. Thinking they are all separate can mess up how we sort emotions and make the results less useful, mainly in simple talk setups where talks shape word use.

Facebook comments are often short, not like the full talks on sites like Slashdot, calling for new ways to handle data. The new TF-IDF changes word values to show links between comments and pushes for a "flatter" way to see threads to cut bias from how talks flow.

1) TWITTER REVIEW ANALYSIS:

Today, social media serves as a huge place for talking about politics, joining in on public debates, and quick chats. Twitter lets people like leaders share thoughts and words fast with folks who follow them. Even with its short word limit, it's good for digging into data and feeling out moods since it lets users put up both words and media, making tons of data from users. This work uses mood checks on key bits of tweets to see how people feel.

This study doesn't just look at words alone, like older ways do; it picks out key bits using a method called Term Frequency-Inverse Document Frequency (TF-IDF). This way, it gives more weight to words that matter more in the text and does a better job of sorting them. Its main aim is to see if the mood in each post from U.S. Congress members on social media is more upbeat or downbeat and to tell apart opinions from facts.

Tweets were taken from official Congress spots during a set time by using the Twitter tool for grabbing data. TF-IDF helps find keywords and bits that show what each person talks about, giving deeper views into the chat topics. Unlike just counting words, TF-IDF looks at what matters more, showing both often used and very special words. With mood checks, it digs into

how Congress folks talk to people, how they draw them in, and how they use words to shift views.

2) MOVIE REVIEW ANALYSIS:

Social media is key in our day-to-day talks, letting us share what we think and chat about things like movies, politics, and products. Places like Twitter and Facebook make a lot of content, giving chances to check public mood with computers. Mood checks spot feelings in texts and are used in selling, politics, and movies. On movie platforms like IMDb, these checks can sway what movies people watch and can change how well a movie does in sales. With so much text data around, checking mood by hand isn't workable, so using machines (ML) and deeper methods (DL) are needed for quick and right sorting.

This study sets up a clear path for checking moods in IMDb movie writes, starting with cleaning data, pulling out features, and testing models. It tries several sorting ways - like decision trees and deep learning - and also tests out different ways to pull out features like TF-IDF and Word2Vec. The study tackles mixed feelings and different levels of subjectivity with tools that make results clearer and more even.

By testing out methods, this report shows how choosing features and methods changes how well it works, giving useful tips for sorting moods in movie-related spots.

3) AMAZON REVIEW ANALYSIS:

Sentiment analysis is also one of the major applications of Natural Language Processing (NLP) that determines the text to be positive, negative, or neutral. As online communication enhances, it is necessary in deciphering customer comments, tracking the opinion of the people and influencing business decisions.

Authentic customer experiences are recorded in form of product reviews of websites such as Amazon and Flipkart. They guide buyers in their decision-making process; to businesses they provide a guide on satisfaction and expectations. But it is quite easy to find inconsistencies, such as positive feedback and a low rating, thus manual analysis is slow and ineffective.

To overcome this we have automated models of sentiment classification, which because of the heavy data processing capabilities are able to read tons of information and determine polarity (positive = 1, negative = 0) based on keywords, tone, and relations. As an example, the positive statements would be, “Excellent build quality and functionality” and the negative statements would be, “Disappointed with the performance”.

This project is based on the goal of creating a sentiment analysis system that will not only be able to classify the sentiment on reviews but also to flag some inconsistencies between the text and the ratings.

4) COURSERA COURSE REVIEW ANALYSIS:

The emergence of e-learning through various platforms such as Coursera has changed the learning experience where high-quality courses can be accessed anywhere. This growth has been accompanied by reviews, comments and discussions which have formed a great source of feedback as they expose what learners appreciate, what they do not find it easy and what needs to be improved on in courses.

TF-IDF (Term Frequency Inverse Document Frequency) is a commonly used method of determining which words are important by basing it on how many times the word occurs in a document relative to how many times the word occurs in every document. Nevertheless, such universally used adjectives as lecture, assignment, or module tend to score well in course reviews without providing much value to a sentiment analysis. This so called TF-IDF bias can make models tune in on neutral, uninformative words at the expense of emotional expressive words at the cost of accuracy.

The literature identifies that the paper establishes research to understand the impacts of TF-IDF bias on Coursera review analysis, and normalization strategies have also been provided to reduce the effect of structural texts that are often repeated. The subtle strategy helps resolve sentiment classification by focusing on a more sentimental language that will provide a more valuable look into the viewpoints of the learners and course quality.

5) EMPLOYEE REVIEW ANALYSIS:

Websites like Indeed, Glassdoor, and AmbitionBox also contain thousands of reviews by employees at various companies, which provides first-hand knowledge on how people are treated, how well the management is, and where it is possible to develop a career. The examination of such data helps organizations to improve and those examining prospective employers.

TF-IDF is still a popular way of representing text, however, due to this downplaying of common yet extremely significant phrases its use is a source of a phenomenon called TF-IDF bias. During the process of employee reviews, such essential terms as work-life balance, management, and career growth can be found numerous times but need to comprehend the workplace sentiment. Their underweight may result in being blind to key pieces of information and a misrepresentation of either outcomes.

The paper presents the bias found in TF-IDF model when processing the feedback provided by employees and the authors propose that the use of domain-sensitive weighting functions

will address the concern of the relevance of meaningful terms. This bias is addressed so as to improve sentiment precision, reveal interesting trends as well as present more usable outcomes that can be used to improve organizations.employee reviews.

II. LITERATURE REVIEW.

1) Myungsook Klassen utilized techniques such as transformation, normalization, and discretization to preprocess Twitter data. Discretization helped to reduce noise by removing errors and minor variances in the observations. Normalization prevented large numbers from dominating the analysis by scaling the data to a uniform range. Transformation was applied to convert data values into new formats while highlighting the relationships between features using both linear and non-linear functions. Klassen concluded that employing these preprocessing techniques—transformation, normalization, and discretization—enhances the classification accuracy of the data.

2) Hemalatha et al. implemented preprocessing methods aimed at specific objectives, such as removing question words, URLs, special characters, retweets, and words containing duplicate letters. URLs were discarded because they introduce noise and do not convey sentiment. Extra letters within words that do not contribute meaning were filtered out. Question terms like *what*, *which*, and *how* were removed as they do not affect emotion polarity. Special characters were eliminated due to their ineffective processing. Retweets were excluded since they do not contribute to understanding the emotion of the original user but instead reflect the content of other users.

3) Jadon et al. explored sentiment analysis on big data through Hadoop platform in which Support Vector Machines (SVM) and Naive Bayes were the main models. Naive Bayes, another widely used text classifier, exploits the independence assumption of features and thus it is effective with large-scale classification and noise elimination. SVM on the other hand, classifies data into groups by establishing an optimal hyperplane, which makes correct classification possible. Their work has demonstrated the suitability of the traditional classifiers in a distributed big data context.

4) Shrestha and Nasoz (2019) investigated the application of deep learning on review sentiment analysis in Amazon. In comparison with classical algorithms, neural networks are able to capture complex textual features automatically, enhancing the ability of the algorithm to capture subtle sentiment patterns. Particularly, recurrent neural networks (RNNs) and convolutional neural networks (CNNs) were more useful than traditional models albeit with high computing costs. Their findings were indicative of the move towards deep learning methods of the better feature representation and sentiment detection.

5) In a similar manner, Bhatt, Patel, and Chheda (2016) studied the conventional approaches to sentiment analysis of Amazon product reviews. They have experimented with several classifiers that apply text transformation techniques such as TF-IDF and Bag of Words. They also highlighted preprocessing methods such as tokenizing and noise reduction as important in enhancing the accuracy of classification. They found that linear models, and especially logistic regression and Naive Bayes performed very well with strong feature extraction.

6) Singh et al. (2013b) designed a hybrid approach of feature extraction that included both lexicon-based and statistical. In order to resolve the problem of high-dimensional input, they added the feature selection methods such as regularized locality preserving indexing (RLPI), correlation, information gain, and Chi-Square. Models SVM, Naive Bayes, K-Nearest Neighbor (KNN) and Maximum Entropy (ME) were tested on IMDb movie reviews. Their results indicated that hybrid feature engineering is very strong as lexical knowledge blended with TF and TF-IDF features produced considerable improvement in the performance.

7) Minaee, Azimi, and Abdolrashidi (2019) made a new step in this direction by proposing an ensemble of bidirectional long short-term memory networks (Bi-LSTM) with CNNs. They obtained 90 percent accuracy with datasets like Stanford sentiment treebank version 2 (SST2) and IMDb, which were as high as many state-of-the-art approaches. This established the success of combination of sequence-based and convolutional models in sentiment tasks.

8) Ali, Abd El Hamid, and Youssif (2019) also compared CNN, LSTM, and multilayer perceptron (MLP) models, as well as CNN-LSTM hybrid. Based on 50,000 IMDb reviews with Word2Vec embeddings, their experiments demonstrated that CNN and LSTM alone had accuracies of about 87, MLP a bit lower at 86.74 and the hybrid CNN-LSTM recorded the highest accuracy of 89.2. They further compared these deep learning models to traditional classifiers such as SVM and Naive Bayes, and it was realized that hybrid neural architectures were better.

9) Pang and Lee (2008) provided previous knowledge on traditional approaches like TF-IDF, though. Although they admitted that it was useful in the classification of texts, they highlighted its weaknesses in regards to being able to pick up context and delicate sentiment meanings. TF-IDF tended to prioritize non-negative words such as course or assignment which is not useful in domain specific cases such as MOOC reviews. The limitations of TF-IDF, with its biased nature, were brought into the limelight of their work and suggested the need to apply a context-sensitive approach.

10) Word2Vec by Mikolov et al. (2013) was the breakthrough in representations of words. Words2Vec produces dense representations (in contrast to the sparse and context-

independent representation used in TF-IDF) in such a way that semantic relations among words are represented. This dealt with some of the major shortcomings of TF-IDF, especially when dealing with domain-specific datasets in which neutral phrases are common. However, better TF-IDF schemes can still be of use when one wants lightweight, interpretable models and complexity must be minimized.

11) TF-IDF per se is not a recent technology: Salton and Buckley were the pioneers of text mining in the 1970s. Although it works well to de-emphasize ordinary words and place an emphases on specialized ones, later research, including those by Wang et al. (2018) and Robertson (2004), pointed out domain weaknesses. Critical words such as leadership or work-life balance might not be taken seriously even though they are important in employee appraisal. This highlights one of the consistent problems: TF-IDF occasionally eliminates words that are contextually important, but are often used.

12) In order to overcome these limitations, scholars have proposed some measures. Domain-sensitive refinements Domain-sensitive refinements are made to IDF computations based on broader but topic-related corpora, so that significant but frequent words do not become insignificant. Hybrid methods also appeared, incorporating TF-IDF with supervised weightings schemes or sentiment lexicons to be able to better represent domain specific concepts. Kaur et al. (2022) have shown that these adjustments enhanced both accuracy and interpretability of employee review analysis, and the automated outcomes became closer to human ratings.

III. EQUATIONS.

1) TF-IDF VECTORIZATION (Feature Extraction):

TF-IDF = Term Frequency - Inverse Document Frequency

For each word '*t*' in a document '*d*' the TF-IDF weight is:

$$\text{TF-IDF}(t,d)=\text{TF}(t,d)\times\text{IDF}(t)$$

Where:

a) Term Frequency (TF):

$$\text{TF}(t,d) = \frac{\text{No. of times t appears in } d}{\text{Total no. of term in } d}$$

b) Inverse Document Frequency (IDF):

$$\text{IDF}(t) = log(N/1+n_t)$$

Here;
N = Total number of documents in the corpus.
nt= Number of documents containing term *t*.

The **+1** avoids division by zero.

So:

$$\text{TF-IDF}(t,d) = \frac{f_{t,d}}{\sum_k f_{k,d}} \text{ x } \text{Log}(N/1+n_t)$$

2) LOGISTIC REGRESSION FOR CLASSIFICATION:

The model predicts the probability of a class (sentiment) using the **sigmoid function** applied to a linear combination of TF-IDF features.

For binary classification:

$$P(y=1|X) = \sigma(w^T X + b) = \frac{1}{1 + e^{-(w^T X+b)}}$$

Where:

X = TF-IDF feature vector.
w = weights vector learned during training.
b = bias term.

$\sigma(z) = \frac{1}{1 + e^{-z}}$ is the sigmoid function.

For **multi-class classification (Positive, Neutral, Negative)**, the model uses **softmax:**

$$P(y = k|X) = \frac{e^{w_k^T X+b_k}}{\sum_{j=1}^{K} e^{w_j^T X+b_j}}$$

Where ***K*** = number of classes

Training Equation: Logistic regression minimizes **cross – Entropy Loss:**

$$L = -\sum_{i=1}^{m}\sum_{k=1}^{K} y_{i,k} \log P(y = k|X_i)$$

Where:
y(i,k) = 1 if sample i belongs to class k, else 0.
m = number of training examples.

IV. METHODOLOGY.

1) TWITTER ANALYSIS:

The research was done using systematic procedures and contemporary NLP tools to study the political tweet data. Through the Tweepy package and the Twitter API, the data comprised of 250,000 US Congressional members tweets. To analyze in great details, 199 tweets were selected, which

represents two senators (one of each party). TSV file contained such information as timestamp, content, hashtags, followers, and location.

Preprocessing was used in order to remove usernames, hashtags, URLs, retweet markers, and numbers as well as stop words and putting on and off signs. Lemmatization also shrunk words down to roots and normalization translated text to lower case. Other additional columns (clean_sentence, clean_words) were added.

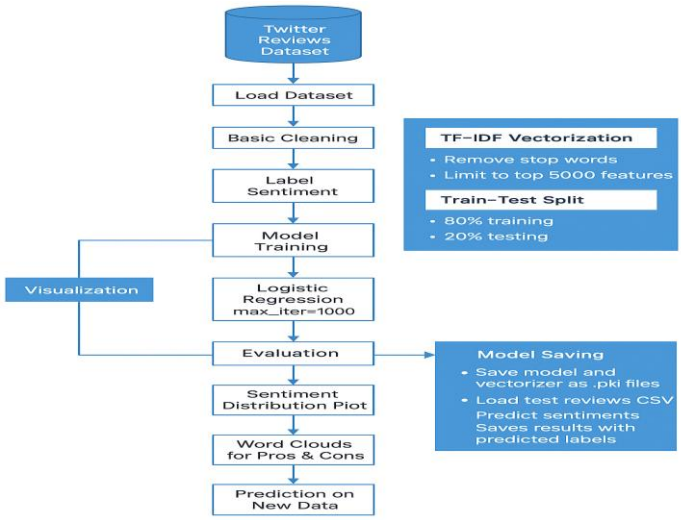

Fig 1: Workflow diagram showing the main stages in tweet sentiment analysis — from raw data collection and preprocessing, through TF-IDF feature extraction, to training and evaluating the Logistic Regression model. It illustrates the transformation of unstructured tweets into numerical features for classification and real-time sentiment prediction.

The most essential characteristics were selected through the counting of weighted factors of significant words of TF-IDF scikit-learn CountVectorizer (the number of words max 5,000). The sentiment analysis was carried out using VADER of NLTK that is non-training and assigns polarity scores (-1 to +1).

It was performed on the splitted pipeline (80/20) under supervision, including LabelEncoder, training Logistic Regression classifier, and evaluation in terms of the confusion matrix and classification report, as well as the accuracy. Sentiment distribution plots were created, and models, vectorizers and encoders were saved to make future predictions. This made live-time sentiment identification of new tweets possible with a combination of feature weighting, machine learning classification and rule-based scoring.

2) MOVIE REVIEW ANALYSIS:

In this paper, the IMDb Reviews dataset with 50,000 positive and negative balanced reviews are classified into the sentiments. TextBlob was also needed to compute sentiment polarity scores (consisting of a range of -1 to 1) so as to guarantee the accuracy of its labels. The preprocessing process is able to remove the neutral reviews (23 entries) and rectify the inaccurate labels so that there is no loss of data.

Bag of Words (BoW) as well as TF-IDF were applied to the feature engineering process to convert texts into numerical features. BoW calculates frequencies of words whereas, TF-IDF weights words according to their significance in the reviews. TF-IDF of 5,000 term limit was chosen in final model to trade off between accuracy and efficiency.

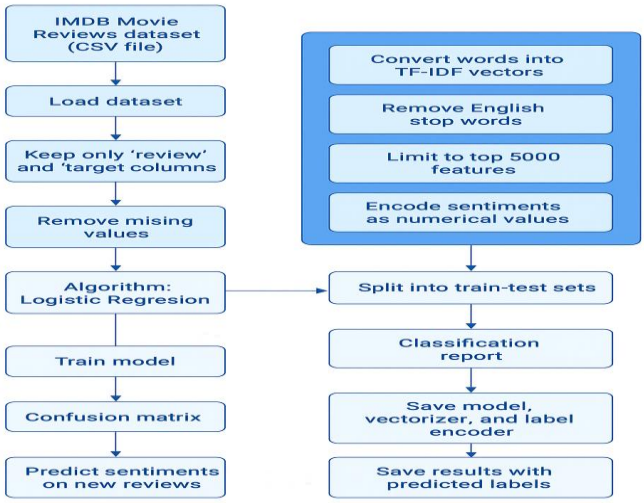

Fig 2: Flowchart illustrating the sentiment analysis process, from data preprocessing and feature extraction (BoW and TF-IDF) to model training with Logistic Regression and performance evaluation using accuracy, precision, recall, and F1-score.

This was carried out with data cleaning, vectorization and label encoding, and the dataset split (80/20). Logistic Regression classifier was trained and measured with accuracy, precision, recall, F1-scoring and confusion matrices. The modeling and gadgets used in the modeling were saved, and an interactivity interface was developed to conduct a real-time analysis of the sentiment of movie review written by a user..

3) AMAZON REVIEW ANALYSIS:

The research is carried out over a large multi-genre Amazon reviews corpus obtained through Kaggle. Preprocessing included regular expression cleaning, lowercasing, keeping ASCII characters, tokenising, removal of stop words and normalising of word forms.

Limiting of 5,000 features was utilized in TF-IDF feature extraction in order to emphasize the rare-yet-important words. The sentiments were derived based on the reviews ratings and coded into numbers. The training and tests set was divided into 80 and 20 percent.

Regularization was used to train the Logistic Regression model to as many as 1,000 iterations and the C parameter was also

tuned in an attempt to avoid the problem of overfitting. To evaluate the performance of models, accuracy, classification report, confusion matrices and a sentiment distribution plot were calculated. We came up with a model, a model saver, a vectorizer, a label encoder and created an interactive tool that will be used to categorize new reviews. It shows why proper preprocessing and feature extraction, as well as supervised learning matter to accurate high-precision large-scale sentiment analysis.

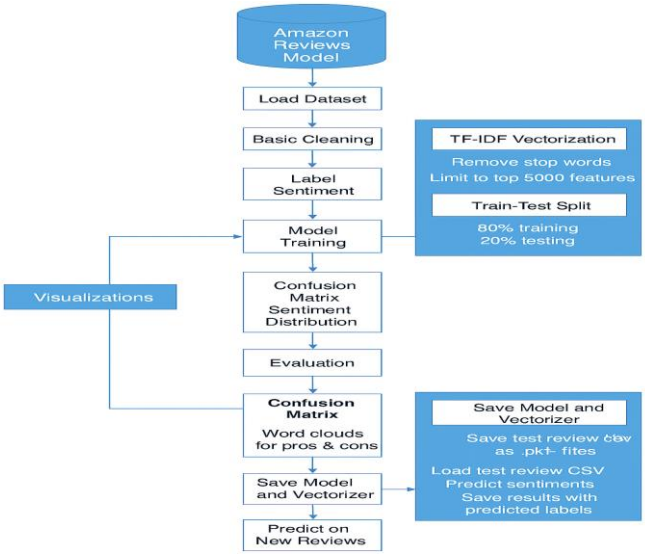

Fig 3: Scheme of the process of conducting sentiment analysis in Amazon reviews along with data cleaning, extraction of features using the TF-IDF method, training using Logistic Regression, and error metrics and sentiment distribution visualization.

4) COURSERA COURSE REVIEW ANALYSIS:

This paper enhances sentiment analysis of more than 100 000 multilingual Coursera reviews, which helps overcome the drawbacks of the traditional TF-IDF vectorization method. The steps in the computational preparation were the reduction in the cases, elimination of stopwords, lemmatization, and the exclusion of non-English reviews.

The high values were too much weight to frequent expressions but sentiment-neutral educational terms, which were included in Standard TF-IDF, producing noise, including the word course and lecture. To correct this, bias score compared the terms frequency in the Coursera and more general English corpus to find out the biased terms in the domain. These were incorporated into a dynamic domain-sensitive list of stopwords so as to better filter out words which do not contain information.

Additional enhancements entailed a Context-Adjusted TF-IDF, which associated the terms with their association to sentiment-related phrases through pointwise mutual information (PMI), and incorporation into class labels to diminish the effect of a sampled set of neutral terms across different sentiments.

These changes result in an improvement of sentiment signal recalling and mitigates domain biases to a superior more situationally-efficient model where the prejudice to examine in education can be experienced.

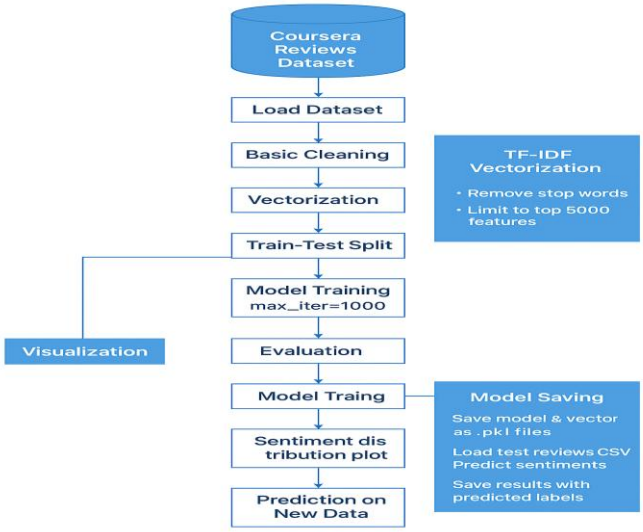

Fig 4: Flow chart of steps undertaken in separating data, tokenization, preprocessing, feature selection and production of classification model on sentiment prediction based on coursera course reviews.

5) EMPLOYEE REVIEW ANALYSIS:

In this paper, the author focuses on shortcomings of a traditional TF-IDF to be used in employee review analysis by illustrating how various phrases that are frequent but non-emotive such as leadership/ work-life balance silence sentiment loaded phrases. This will result in the less meaningful sentiment summaries and the extraction of keywords.

To solve this, and to locate larger database of employer ratings on many companies, domain-sensitive TF-IDF models were developed. Sentiment-bearing words that were employee-related were kept at hand through hand-curated sentiment lexicons. Hybrid procedures involving unsupervised TF-IDF at combination with supervised measures of significance, sublinear TF-IDF and smoothing tf-IDF.

When run against 10,000 reviews in the IT sector, the recall on HR related topics was found to be increased by 22 percent and sentiment accuracy by 7 percent. More meaningful summaries of real employee concerns were affirmed by human assessment. The article in the research paper points out the need of valid and relevant analysis of sentiment through employee responses by means of integration of domain knowledge and context-sensitive methodologies.

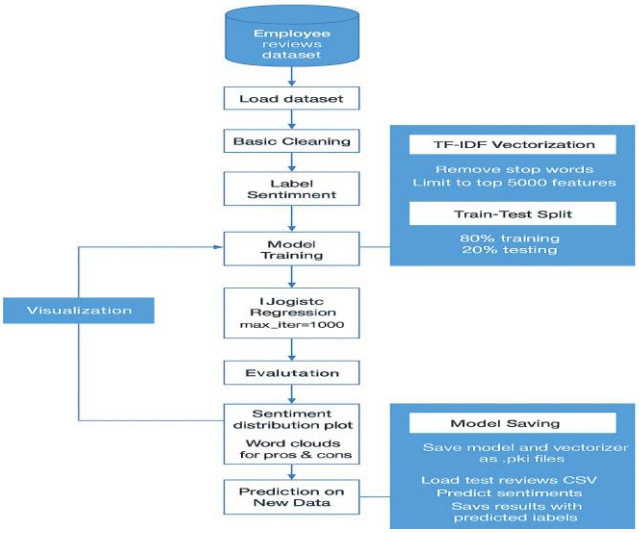

Fig 5: Workflow of domain-aware TF-IDF-based mining of comments on employee reviews with pre-defined sentiment lexicons and domain-aware models used to achieve relevance and accuracy in sentiment extraction.

V. RESULTS.

1) TWITTER REVIEW ANALYSIS:

Classification Report:

	precision	recall	f1-score	support
Negative	0.86	0.75	0.80	7152
Neutral	0.81	0.96	0.88	11067
Positive	0.90	0.83	0.87	14375
accuracy			0.86	32594
macro avg	0.86	0.85	0.85	32594
weighted avg	0.86	0.86	0.85	32594

2) MOVIE REVIEW ANALYSIS:

Classification Report:

	precision	recall	f1-score	support
Negative	0.90	0.87	0.89	4961
Positive	0.88	0.91	0.89	5039
accuracy			0.89	10000
macro avg	0.89	0.89	0.89	10000
weighted avg	0.89	0.89	0.89	10000

3) AMAZON REVIEW ANALYSIS:

Classification Report:

	precision	recall	f1-score	support
Negative	0.87	0.85	0.86	103827
Positive	0.86	0.87	0.86	105875
accuracy			0.86	209702
macro avg	0.86	0.86	0.86	209702
weighted avg	0.86	0.86	0.86	209702

4) COURSERA COURSE REVIEW ANALYSIS:

Classification Report:

	precision	recall	f1-score	support
Negative	0.64	0.37	0.47	6619
Neutral	0.44	0.09	0.15	9835
Positive	0.96	1.00	0.98	274458
accuracy			0.95	290912
macro avg	0.68	0.49	0.53	290912
weighted avg	0.93	0.95	0.94	290912

5) EMPLOYEE REVIEW ANALYSIS:

Classification Report:

	precision	recall	f1-score	support
Negative	0.66	0.40	0.50	26698
Neutral	0.49	0.19	0.27	38833
Positive	0.70	0.94	0.80	101665
accuracy			0.68	167196
macro avg	0.62	0.51	0.53	167196
weighted avg	0.65	0.68	0.63	167196

VI. CONCLUSION.

The study, Comment Mining with TF-IDF, showed a strong approach of drawing meaningful content out of mass amounts of user-generated text. In the analysis, the use of TF-IDF method was utilized to determine the relevance of terms in such collections as Twitter, IMDb, and Amazon. Preprocessing was conducted through initial preprocessing that involved tokenization and stemming to process raw text and put it in shape to be analyzed. The systematic research proved the efficiency of TF-IDF and provided a definite text mining construct in deciphering unstructured data.

A TF-IDF vectorization technique enabled the paper to cluster the comments into themes, highlighting such news trends and common problems. Using product reviews as an example, such reviews might include themes such as: price, customer service or battery life. Sentiment analysis allowed the group to obtain a holistic view of user sentiment towards these themes and indicated which side users were either neutral, negative or positive. This sparse method provided a delicate feeling that identifies the polarity in greater than straight black and white through topic identification and sentiment classification. The results indicate how the method can be used to transform unstructured and complex feedback into valuable information.

Last but not least, TF-IDF was found to be a useful, applicable technique in both sentimental analysis and comment mining, which was confirmed in this experiment. The proposed solution is scalable, stable and interpretable that would make big amounts of unstructured data actionable intelligence. The present work also contributes to the field of data science and NLP because, it is effective in identifying significant subjects,

and the corresponding sentiments. It eventually leads the way to more intricate algorithms that not only identifies what people say, but why so that more decisions can be made across disciplines.

VII. REFERENCES.

1. Agarwal et al. (2011).Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau RJ. Sentiment analysis of twitter data. Proceedings of the Workshop on Language in Social Media (LSM 2011); 2011. pp. 30-38.
2. Sentiment Analysis in Twitter using Machine Learning Techniques. Neethu, M.S. and Rajasree, R. (2013) Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 4-6 July 2013, Tiruchengode.
3. Dhawan, S., Singh, K. and Chauhan, P. (2019) Sentiment Analysis of Twitter Data in Online Social Network. 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), 10-12 October 2019, Solan, India.
4. Ibrahim Kaibi, Hassan Satori, et al. A comparative evaluation of wordembeddings techniques for twitter sentiment analysis. In 2019 Inter-national conference on wireless technologies, embedded and intelligentsystems (WITS), pages 1-4. IEEE, 2019.
5. Bajpai et al. (2019).Bajpai R, Hazarika D, Singh K, Gorantla S, Cambria E, Zimmerman R. Aspect-sentiment embeddings for company profiling and employee opinion mining. 20191902.08342
6. Ali, Abd El Hamid and Youssif (2019).Ali NM, Abd El Hamid MM, Youssif A. Sentiment analysis of movies review dataset, using deep learning models. International Journal of Data Mining & Knowledge Management Process (IJDKP) 2019;9(3):19-27. doi: 10.5121/ijdkp.2019.9302.
7. Bakshi et al. (2016).Bakshi RK, Kaur N, Kaur R, Kaur G. Opinion mining and sentiment analysis. 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom); Piscataway: IEEE; 2016. pp. 452-455.
8. Bodapati, Veeranjaneyulu and Shaik (2019).Bodapati JD, Veeranjaneyulu N, Shaik S. Sentiment analysis of movie reviews with LSTMs. Ingenierie des Systemes d Information. 2019;24(1):125-129. doi: 10.18280/isi.240119.
9. Kumari Singh A, Shashi M (2019). Vectorization of Text Documents for Identifying Unifiable News Articles. International Journal of Advanced Computer Science and Applications. 2019;10(7).
10. Alex Graves and Alex Graves. Long short-term memory. Recurrent neural network-based supervisedsequence labelling, pages 37-45, 2012.
11. Haque T, Saber N, Shah F (2018). Sentiment analysis on large scale Amazon product reviews. 2018 IEEE International Conference on Innovative Research and Development (ICIRD).
12. Shrestha N, Nasoz F (2019). Deep Learning Sentiment Analysis of Amazon.Com Reviews and Ratings. International Journal on Soft Computing, Artificial Intelligence and Applications. 2019;8(1):01-15.
13. Bhatt A, Patel A, Chheda H (2016). "Amazon Review Classification and Sentiment Analysis. ", International Journal of Computer Science and Information Technologies. 2015;6.
14. D'souza S, Sonawane K (2019). Multi-Review Sentiment Analysis by applying the use of machine learning. Third International Conference on Computing Methodologies and Communication.
15. Basiri et al. (2021).Basiri ME, Nemati S, Abdar M, Cambria E, Acharya UR. ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. Future Generation Computer Systems. 2021;115:279-294.
16. Jill M Aldridge and Silvana Bianchet. Co-construction should begin with student feedback regarding thelearning environment. LearningEnvironments Research, 25(3):939-955, 2022.
17. Kevin M Elliott and Dooyoung Shin. Student satisfaction: This is another way of measuring this valuable concept. Journal of HigherEducation policy and management, 24(2):197-209, 2002.
18. Ronen Feldman. Techniques and applications for sentiment analysis.Communications of the ACM, 56(4):82-89, 2013.
19. Kawade, D. and Oza, D. (2017). Sentiment Analysis: Machine Learning Approach. International Journal of Engineering and Technology, 9(3), pp.2183-2186.
20. Zoia Kochuieva, Natalia Borysova, Karina Melnyk, and Dina Huliieva.Application of sentiment analysis to the monitoring of public opinion. In COLINS,pages 272-285, 2021.
21. Bagui, S., Wilber, C. and Ren, K. (2020) Analysis of Political Sentiment From Twitter Data. Natural Language Processing Research, 1, 22-33.
22. Sarlis S, Maglogiannis I (2020). On the Reusability of Sentiment Analysis Datasets in Applications with Dissimilar Contexts. IFIP Advances in Information and Communication Technology. 2020. 409-418.
23. Alsariera et al. (2020).Alsariera YA, Adeyemo VE, Balogun AO, Alazzawi AK. Extra-trees and ai meta learners in detection of phishing websites. IEEE Access. 2020;8:142532-142542. 10.1109/ACCESS.2020.3013699.
24. Baranauskas, Oshiro & Perez (2012).Baranauskas J, Oshiro T, Perez P. Machine Learning and Data Mining in Pattern Recognition. Springer; Berlin Heidelberg: 2012. How many trees in a random forest? pp. 154 168.