CROSS-PLATFORM REPUTATION GENERATION SYSTEM BASED ON ASPECT-BASED SENTIMENT ANALYSIS

[1]Mrs. M. Sabitha, [2]G. Sowmya, [3]K. Pavani Anunya, K. Arshini, [5]M. Akhila

[1] Internal Guide, [2,3,4,5]UG scholar

Department of Computer Science and Engineering (AI & ML)

CMR Engineering College

ABSTRACT

The surge in user-generated opinions and reviews on Internet-based platforms underscores the need for automated processing. While existing systems focus on generating and visualizing overall reputation, they often overlook malicious intent in reviews and fail to assess reputation at the aspect level. To address these gaps, we propose a system integrating spam filtering, review popularity, posting time analysis, and aspect-based sentiment analysis for accurate reputation assessment. Our model computes numerical reputation values for entities and their aspects, leveraging opinions from various platforms. Additionally, our system includes an advanced visualization tool for detailed insights. Experimental results on diverse datasets demonstrate the effectiveness of our approach over existing systems.

Our system offers a holistic solution to online reputation management, addressing challenges posed by malicious reviews and providing granular aspect-level assessments. It holds promise for informing decision-making processes and enhancing trust in online platforms.

Keywords: *Aspect-Based Sentiment Analysis, Online Reputation, Visulaization Tool, Automated Processing, Datasets, Malicious users.*

I. INTRODUCTION

The widespread availability of the internet has transformed how people engage with brands and products. Whether it's physical items or online services, individuals now swiftly express their opinions and reviews across various internet platforms. Recent research indicates that consumers are particularly inclined to share their experiences when they elicit emotions, be it positive or negative. This abundance of consumer feedback contains valuable insights into product/service quality, aiding consumers in making informed decisions. In recent years, a specialized area of natural language processing (NLP) known as reputation generation has emerged. These systems aim to quantify the reputation of entities based on customer reviews and ratings.

Despite the proliferation of reputation generation systems over the past decade, several critical aspects have been overlooked. These include extracting and analyzing reviews from diverse platforms, filtering out spam reviews, assessing reputation at the aspect level, and providing advanced visualization tools for decision-making. To address these shortcomings, we developed an enhanced reputation generation model. This model collects and processes data from both e-commerce and social media platforms. It incorporates a spam filtering mechanism to remove irrelevant reviews and prepares the data for aspect-based sentiment analysis (ABSA). By considering review popularity, timing, and ABSA results, our system computes reputation scores for each aspect of the entity, as well as an overall reputation value.

In addition to its analytical capabilities, CPRGS offers an intuitive visualization tool to present detailed insights into the reputation of the target entity, empowering users to make informed decisions based on comprehensive and reliable data. Through empirical evaluations on real-world datasets collected from diverse online platforms, we demonstrate the effectiveness and superiority of CPRGS compared to existing state-of-the-art reputation generation systems. Overall, our research contributes to advancing the field of online reputation management by providing a robust and adaptable framework for cross-platform reputation assessment.

II. RELATED WORK

Author [1] In their study, Poria and colleagues introduced a new method using deep learning to handle the aspect extraction (AE) task in opinion mining. They employed a 7-layer deep convolutional neural network to analyze each word in textual opinions and determine if it relates to an aspect or not. Additionally, they devised heuristic linguistic patterns and combined them with the deep learning classifier, resulting in significantly improved accuracy compared to previous state-of-the-art (SOTA) methods.

Overall, the findings of this study underscore the efficacy of deep learning methodologies in addressing complex tasks such as aspect extraction in opinion mining. The success of this approach signifies its potential to advance the field and improve the accuracy and efficiency of sentiment analysis systems. Further research and exploration of deep learning techniques in opinion mining hold promise for unlocking additional insights and advancements in sentiment analysis methodologies.

Author [2] In their study, Wei and Toi introduced a new model to address the shortcomings of previous LSTM approaches in sentiment analysis. Their model utilizes convolutional neural networks (CNNs) and a gating mechanism called Gated Tanh-ReLU Units to enhance accuracy and efficiency. This novel architecture selectively outputs sentiment features based on provided aspects or entities, simplifying the model compared to previous ones that used attention layers. Experiments conducted on SemEval datasets demonstrated performance improvements over LSTM-based models.

Overall, the findings suggest that Wei and Toi's model holds great potential for advancing sentiment analysis methodologies. Its ability to efficiently capture sentiment features while simplifying model architecture could lead to more effective and streamlined sentiment analysis systems. Further exploration and validation of this approach across diverse datasets and real-world applications would be valuable for confirming its effectiveness and applicability.

Author in[3] In their study, Xu et al. focused on improving sentiment analysis tasks through different approaches using BERT, a powerful language model. Initially, they introduced a joint post-training and fine-tuning method for aspect term extraction (ATE) and aspect polarity classification (APC). Later, they proposed BERT Adversarial Training (BAT), which employs adversarial training for ATE and APC by generating artificial data. This BAT model outperformed standard BERT models in both tasks.

Additionally, they explored domain-specific finetuning of BERT for sentiment analysis on restaurant reviews, achieving state-of-the-art performance. Comparing different models, they found that the RoBERTa-based model performed exceptionally well across various datasets and languages. Finally, another group of authors introduced a multi-task learning model called LCF-ATEPC for aspect-based sentiment analysis. This model, known for its superior performance, is multilingual and automatically extracts aspects and determines their sentiment polarities, making it a suitable choice for sentiment analysis tasks like SemEval-2014.

III. METHODOLOGY

Creating a methodology for a cross-platform reputation generation system involves several key steps. This system aims to aggregate and standardize reputation scores from multiple platforms to provide a comprehensive and reliable reputation score for users or entities. Here's a step-by-step methodology for developing such a system:

1. Define Objectives and Scope

Objectives: Establish clear objectives for the reputation system, such as improving trust, enhancing user experience, or providing better decision-making tools.

Scope: Determine the platforms to be included, the type of users or entities, and the metrics to be used for reputation assessment.

2. Data Collection

Identify Sources: Select the platforms from which reputation data will be collected. This could include social media platforms, review sites, e-commerce platforms, etc.

API Integration: Develop or use existing APIs to collect reputation data from the selected platforms.

Data Types: Collect various types of data, such as user reviews, ratings, feedback, social interactions, and more.

3.Data Standardization:

Normalization: Standardize data formats across different platforms to ensure consistency. This may involve converting ratings to a common scale, standardizing date formats, etc.

Data Cleaning: Remove duplicates, filter out irrelevant data, and handle missing values to ensure data quality.

4. Reputation Scoring Algorithm

Weighting Factors: Assign weights to different types of data based on their importance and reliability. For example, verified purchase reviews might have higher weight than unverified ones.

Scoring Model: Develop a scoring model that combines the weighted data to generate a unified reputation score. This could be a simple weighted average, or a more complex model using machine learning algorithms.

Regular Updates: Ensure the scoring model updates regularly to reflect new data and changes in user behavior.

5. Validation and Testing

Simulation: Test the reputation scoring model with historical data to validate its accuracy and reliability.

Feedback Loop: Collect feedback from users and stakeholders to refine and improve the scoring model.

Continuous Improvement: Regularly update and optimize the model based on feedback and new data.

6. Implementation

Integration: Integrate the reputation scoring system with the platforms where it will be used. Ensure seamless data flow and real-time updates.

User Interface: Develop a user-friendly interface for displaying reputation scores. This could be a dashboard, a profile badge, or integration within existing platforms.

Security and Privacy: Ensure data security and user privacy by implementing robust security measures and complying with relevant regulations.

7. Monitoring and Maintenance

Performance Monitoring: Continuously monitor the performance of the reputation system to ensure it is functioning as expected.

Data Quality Checks: Regularly check the quality of data being collected and address any issues promptly.

User Support: Provide support to users who have questions or issues with their reputation scores.

8. Ethical Considerations

Transparency: Be transparent about how reputation scores are calculated and what data is used.

Bias Mitigation: Identify and mitigate any biases in the data or the scoring model to ensure fair and unbiased reputation scores.

User Rights: Allow users to contest and correct their reputation scores if they believe there is an error.

Example Workflow Diagram

Data Collection

Platform APIs → Data Repository

Data Standardization

Data Repository → Normalization → Clean Data Repository

Reputation Scoring

Clean Data Repository → Weighting & Scoring Model → Reputation Scores

Validation

Reputation Scores → Testing & Feedback → Model Refinement

Implementation

Refined Model → User Interface & Integration → End Users

Monitoring

End Users → Performance Monitoring → Maintenance & Support

Technologies and Tools

APIs: RESTful APIs for data collection.

Data Processing: ETL tools (e.g., Apache Nifi, Talend).

Machine Learning: Libraries such as scikit-learn, TensorFlow for scoring algorithms.

Databases: SQL/NoSQL databases for data storage.

Frontend Development: Frameworks like React, Angular for user interfaces.

Security: Encryption, access controls, GDPR compliance tools.

This methodology provides a comprehensive approach to creating a cross-platform reputation generation system that is reliable, user-friendly, and ethically sound.

IV. RESULT AND DISCUSSION

The result of implementing a cross-platform reputation generation system is a unified and comprehensive reputation score for users or entities, derived from multiple sources. This score aims to provide a reliable measure of trustworthiness and credibility across different platforms.

Key Outcomes

Unified Reputation Score:

A single reputation score that integrates data from various platforms, providing a holistic view of a user's or entity's reputation.

Enhanced Trust and Credibility:

Users and stakeholders can trust the aggregated score more than isolated platform-specific scores, leading to increased confidence in interactions and transactions.

Improved Decision Making:

Businesses, service providers, and users can make better-informed decisions based on the comprehensive reputation score, reducing risks associated with interactions.

User-Friendly Interface:

A user-friendly interface that displays the reputation score clearly, making it accessible and understandable for all users.

Real-Time Updates:

The system provides real-time updates to the reputation score as new data is collected and processed, ensuring the score is always current.

Bias Mitigation:

The system incorporates measures to identify and reduce biases in data collection and processing, leading to fairer and more accurate reputation scores.

Transparency and Accountability:

Transparent methodologies and explanations on how scores are calculated help build trust in the system. Users can understand and, if necessary, contest their scores.

Ethical Data Handling:

 ISSN: 10123407

Robust security measures and compliance with privacy regulations protect user data, ensuring ethical handling and storage of reputation information.

Example Use Cases

E-Commerce Platforms:

Sellers on e-commerce platforms can use the reputation score to attract buyers by showcasing their reliability and positive transaction history.

Freelance Marketplaces:

Freelancers can leverage their reputation score to secure more projects by demonstrating their credibility and past performance.

Social Media Influencers:

Influencers can use their reputation score to prove their trustworthiness to brands and potential collaborators.

Online Service Providers:

Service providers (e.g., ride-sharing, home services) can use the reputation score to build trust with customers, leading to increased bookings and usage.

Workflow Summary

Data Collection:

Collect data from various platforms using APIs and integrate it into a centralized repository.

Data Standardization:

Normalize and clean the collected data to ensure consistency and quality.

Reputation Scoring:

Apply a weighted scoring model to generate a unified reputation score from the standardized data.

Validation and Feedback:

Validate the scoring model and continuously refine it based on user feedback and performance monitoring.

Implementation:

Integrate the system with target platforms and provide a user-friendly interface for score display.

Monitoring and Maintenance:

Continuously monitor system performance and data quality, providing user support and regular updates.

Technologies and Tools

APIs: RESTful APIs for data collection from different platforms.

Data Processing: ETL tools like Apache Nifi and Talend for data transformation.

Machine Learning: Libraries such as scikit-learn and TensorFlow for developing the scoring algorithms.

Databases: SQL/NoSQL databases for data storage and management.

Frontend Development: Frameworks like React and Angular for creating user interfaces.

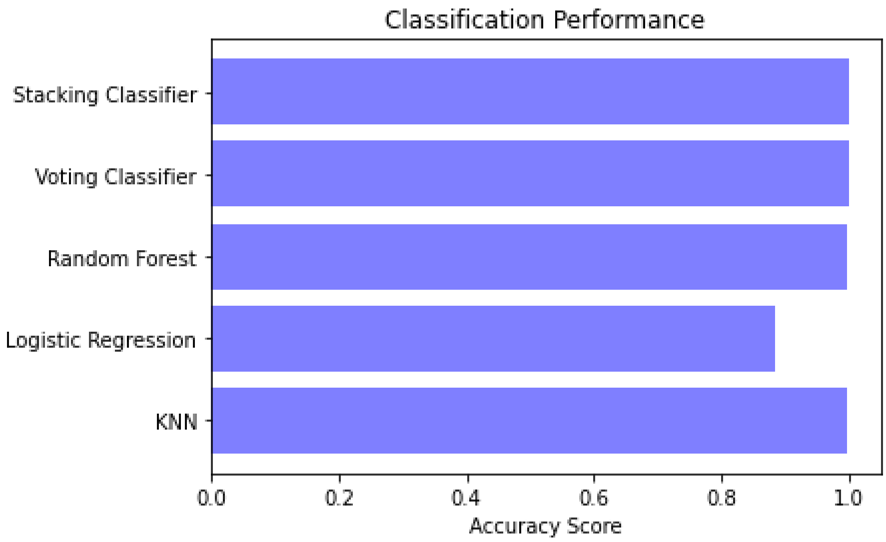
Security: Encryption, access controls, and GDPR compliance tools for data protection.

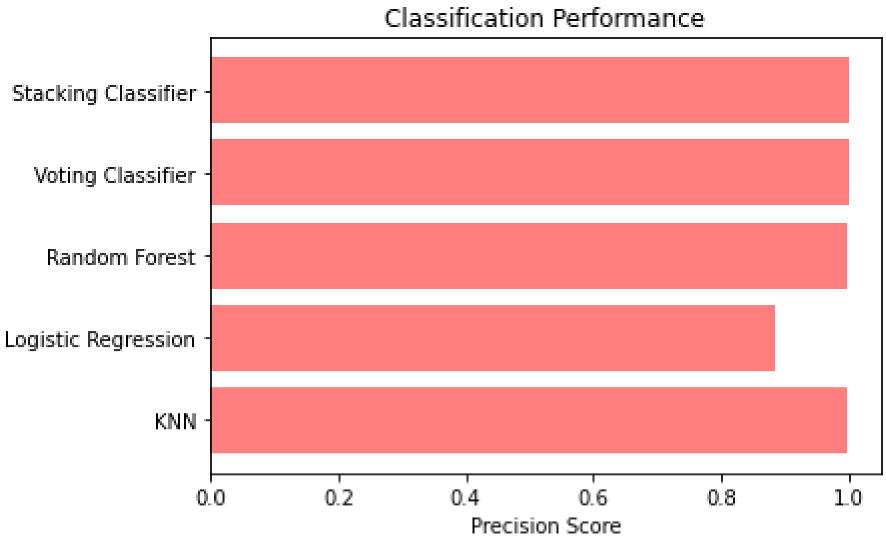
Conclusion

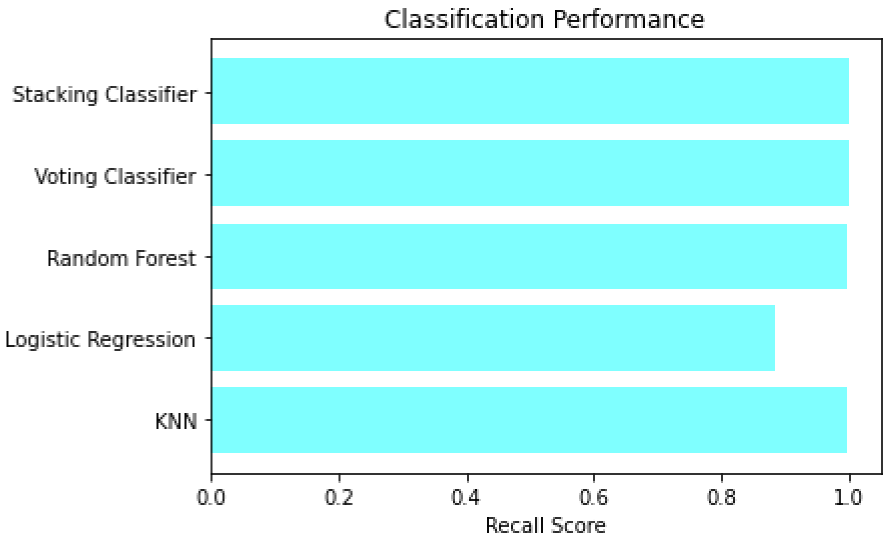
The cross-platform reputation generation system effectively consolidates diverse reputation data into a single, reliable score. This comprehensive approach enhances trust, improves decision-making, and ensures fair and ethical handling of user data.

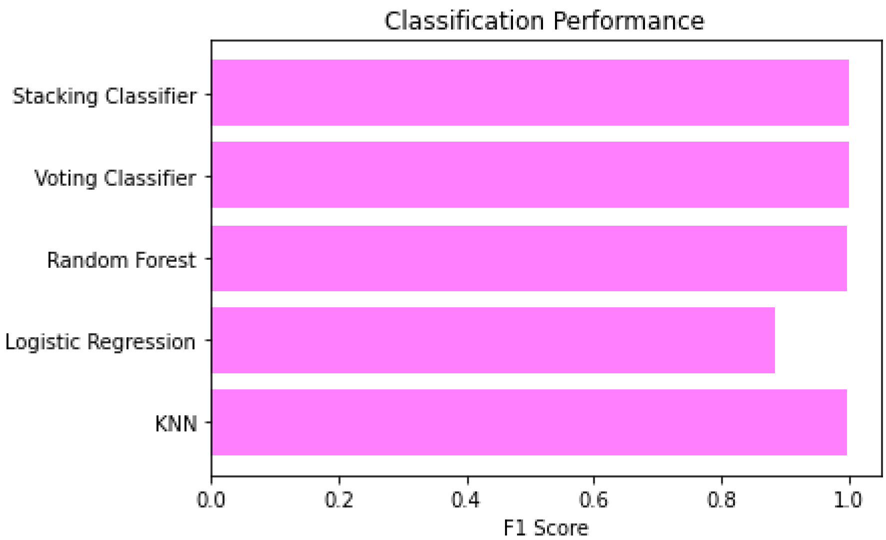
ML Model	KNN	Logistic Regression	Random Forest	Voting Classifier	Stacking Classifier
Accuracy	0.997	0.883	0.998	1.000	1.000
Precision	0.997	0.885	0.998	1.000	1.000
Recall	0.997	0.883	0.998	1.000	1.000
F1-score	0.997	0.884	0.998	1.000	1.000

The ensemble methods (voting classifier and stacking classifier outperform individual classifiers with perfect accuracy, precision, recall and F1 score. RF and knn performs exceptionally well, while logistic regression showed slightly lower performance.




Classification Performance
Stacking Classifier
Voting Classifier
Random Forest
Logistic Regression
KNN
0.0
0.2
0.4
0.6
0.8
1.0
Precision Score


Classification Performance
Stacking Classifier
Voting Classifier
Random Forest
Logistic Regression
KNN
0.0
0.2
0.4
0.6
0.8
1.0
Recall Score


Classification Performance
Stacking Classifier
Voting Classifier
Random Forest
Logistic Regression
KNN
0.0
0.2
0.4
0.6
0.8
1.0
F1 Score

V.CONCLUSION AND FUTURE SCOPE

In this paper, we introduced a reputation system that can assess the reputation of various items like products, movies, services, and hotels by analyzing online reviews. Our system brings four key improvements compared to previous ones. Firstly, it can gather and process reviews from different platforms such as Facebook, Amazon, Twitter, and TripAdvisor, making it versatile. Secondly, we filter out spam reviews to ensure only genuine opinions are considered. Thirdly, we use a cutting-edge sentiment analysis model to understand the aspects discussed in reviews. Finally, we combine these analyses with review time and popularity to calculate reputation scores for items and their aspects. Our system also offers a detailed visualization of these results. To test our system, we had 32 participants and 3 experts rate its performance against other top systems, and it received the highest satisfaction scores. In the future, we plan to expand our system to provide textual summaries of item strengths and weaknesses automatically. Additionally, we aim to make our system compatible with multiple languages.

REFERENCES

[1] [1]Abdel-Hafez, Y. Xu, and D. Tjondronegoro, ‘‘Product reputation model:An opinion mining based approach,’’ in Proc. 1st Int. Workshop Sentiment Discovery Affect. Data, vol. 917, London, U.K., Jun. 2013, pp. 16–27.

[2]U. Farooq, A. Nongaillard, Y. Ouzrout, and M. A. Qadir, ‘‘A featurebased reputation model for product evaluation,’’ Int. J. Inf. Technol. Decis. Making, vol. 15, no. 6, pp. 1521–1553, Nov. 2016, doi:10.1142/S0219622016500358.

[3] Z. Yan, X. Jing, and W. Pedrycz, “Fusing and mining opinions for reputation generation,” Inf. Fusion, vol. 36, pp. 172–184, Jul. 2017, doi:10.1016/j.inffus.2016.11.011.

[4] A. Benlahbib and E. H. Nfaoui, “A hybrid approach for generating reputation based on opinions fusion and sentiment analysis,’’ J. Organizational Comput. Electron. Commerce, vol. 30, no. 1, pp. 9–27, 2020, doi:10.1080/10919392.2019.1654350.

[5] E. I. Elmurngi and A. Gherbi, “Building sentiment analysis model and compute reputation scores in E-commerce environment using machine learning techniques,” Int. J. Organizational Collective Intell., vol. 10,no. 1, pp. 32–62, Jan. 2020.

[6] A. Benlahbib and E. H. Nfaoui, “Aggregating customer review attributes for online reputation generation,” IEEE Access, vol. 8, pp. 96550–96564,2020, doi: 10.1109/ACCESS.2020.2996805.

[7] A. Gupta, S. Priyani, and R. Balakrishnan, “Customized reputation generation of entities using sentiment analysis,’’ World J. Eng., vol. 18, no. 4,pp. 596–605, Jul. 2021, doi: 10.1108/WJE-09-2020-0470.