Grenze International Journal of Engineering and Technology, June Issue

E-Commerce Authenticity: A Deep Learning and Aspect-based Fusion Approach for Fake Review Detection

K Amrutasagar[1], Goli Sri Godha[2], Chidipudi Tejaswini[3] and Chennu Venkata Naga Gopi[4]
[1-4]Seshadri Rao Gudlavalleru Engineering College/CSE (AI & ML), Gudlavalleru, India
Email: sagar.chintu89@gmail.com, srigodha9@gmail.com, chidipudit@gmail.com, venkatanagagopichennu@gmail.com

***Abstract*— Trust in online shopping suffers when false feedback appears. Although influencing customer choices, such content introduces misleading paths through altered opinions. Detection systems respond by identifying mixed forms of deceptive input via thorough examination methods. A structure combining various streams of written traits operates with advanced one-dimensional convolution layers followed by two-way long short-term memory units tailored specifically. Focus increases on key elements thanks to an intelligent marking method. Examine the units of text closely. Preparation occurs first, unfolding in stages without immediate visibility. A mix of Doc2Vec, GloVe, and TF-IDF weighted n-grams forms the base data. Once compressed via SVD, meaning shifts are tracked by grammatical dependencies rather than deep neural stacks. Built upon SBERT representations, emotional context enters by mapping subtle bonds among moods and judgments. Despite differing routes, sequence inputs merge with auxiliary signals, refining awareness of setting during processing. Notably influenced by emotion, interpretation of complex cues grows fivefold in impact. Layered checks operate within the framework, complementing a collective structure shaped by diverse predictors. Once evaluation ends, evidence emerges: the XGBoost fusion stage lifts precision and resilience. In trials, results stand out clearly - truthful, deceitful, favorable, and unfavorable cases are consistently separated. A single investigation examined responses. Following that work, scalability methods entered the design of deep learning systems aimed at improving validation processes within online trade settings.**

***Index Terms*— E-Commerce Authenticity, Fake Review Detection, Deep Learning, Conv1D, BiLSTM, Attention Mechanism, Doc2Vec, TF–IDF, SBERT, Sentiment Analysis, Ensemble Learning, XGBoost, NLP.**

I. INTRODUCTION

In the digital era, e-commerce has become an woven into worldwide commerce, helping shoppers access products from afar. Buying stuff often depends heavily on what shows up in web searches. People share thoughts through reviews, their voices shaping what others see. Though real, what people say can guide others on whether a thing works well. Because it works well yet stays dependable, the quick spread of counterfeit or false feedback has badly weakened the online platforms aren't always honest. Scams pop up where buyers expect safety. People often make up reviews just to boost numbers. Some reviews lift one item while quietly dimming another, leaving a mark that sticks around, misinformation, financial loss, and reduced consumer confidence [13], [14].

Grenze ID: 01.GIJET.12.2.203_24

Traditional methods for detecting deceptive reviews, some early systems used basic algorithms instead of deep learning. Crafted by hand, the way words are shaped matters. Even so, these methods offer starting results, yet they falter to capture deep semantic patterns, emotional odd shifts, yet careful word choices that characterize modern deceptive review generation. The deep learning and natural language progress, processing (NLP) has opened new opportunities for shaping words through deeper forms, making possible learning systems pick up how words connect, what they mean, then feelings behind them, pushed by signals that work better [6], [8].

This study introduces a system called E-Commerce Authenticity that combines deep learning with aspect-focused analysis to spot fake product reviews. Instead of relying on single methods, it pulls together several natural language processing tools to improve accuracy in identifying real versus false feedback. Running behind the scenes is a mix of layered Conv1D units plus bidirectional LSTM networks alongside a unique attention setup [18] - this allows spotting fine-grained word choices while also tracking broader sentence meaning across longer texts. To boost performance further, additional signals like term frequency measures using n-grams, vector outputs from Doc2Vecand GloVe models, contextualized SBERT encodings [11], emotional tones, along with overall sentiment strength feed into the decision process, offering richer clues about dishonest content.

With each test split carefully chosen, five rounds of balanced validation help check how well the method holds up across different kinds of product feedback. Instead of relying on a single model, layers of neural networks feed into an XGBoost stacker, bringing steadier results no matter the review style [9], [10]. Driven by real-world needs, this work targets online shopping sites that need reliable tools to catch fake user opinions at scale. By spotting misleading content faster, systems like these can foster fairer environments where buyers feel more confident about their choices. Trust grows quietly when users sense fewer distortions behind star ratings.

II. Related Work

With time, methods for spotting dishonest reviews have changed, driven by rising numbers of false opinions online. At first, analysts focused on word choices, details about who wrote them, and how reviewers acted. A closer look at deceptive comments began with work directed by Myle Ott [1], [2], uncovering how false ratings often carry exaggerated feelings along with unusual word patterns. While techniques such as Naïve Bayes, Support Vector Machines, and Logistic Regression delivered modest outcomes, detecting finer details in spoken style remained beyond reach. Even with advancement, complete understanding slipped away when those initial models met layered speech traits.

Afterward, progress in learning representations brought forth methods where text appeared as spread-out patterns enhancing meaning capture. Instead of isolated terms, systems began using frameworks like Word2Vec - introduced by Tomas Mikolov [3] - and Doc2Vec from Quoc Le [4], translating words and whole texts into compact numerical forms reflecting surrounding context. Following that, a method named GloVe, shaped by Jeffrey Pennington [5], relied on overall word pairing frequencies across large datasets to generate insightful embeddings. Unlike earlier sparse models based solely on term counts, these vector-based formats offered richer structure through learned associations. Their adoption became common in analyzing fake product evaluations due to stronger descriptive power.

Improvements in deep learning strengthened how detection models perform. Because of their design, Long Short-Term Memory and Bidirectional LSTM structures handled extended dependencies in text well, making it possible to interpret connections throughout full reviews. At the same time, Convolutional Neural Networks showed skill at spotting nearby word arrangements and expressions. With these strengths in mind, combined designs emerged - mixing convolutional and recurrent networks - to make use of both immediate details and sequence-based context when classifying text. While each method contributed uniquely, integration allowed broader pattern recognition without losing sensitivity to structure.

Recently, progress in language processing has come through transformer-driven systems. Instead of simple word matching, these tools interpret meaning based on surrounding context. Among them, BERT and its variants - RoBERTa, for instance - detect nuanced shifts in expression across sentences. Because false statements tend to show slight mismatches in tone or logic, such models excel at spotting irregularities. Deceptive content, especially fabricated product feedback, frequently reveals itself through awkward emphasis or inconsistent framing.

Alongside context awareness, studies frequently examine feelings and moods when spotting false statements. Systems like VADER together with the NRC Emotion Lexicon detect mood direction and mental traces within written words. Deceptive feedback tends to show irregular shifts in emotion or overstated tones - this makes affective markers useful for recognizing dishonest evaluations.

One path worth exploring uses methods where several models work together, boosting reliability through combination. Rather than depend on just one predictor, outputs from neural networks get merged with older algorithm types like XGBoost or Random Forest. Such mixed systems show stronger results when tested on varied data collections.

Following recent progress, this work introduces a mixed-method approach for spotting false feedback, merging diverse feature capture with a layered network involving one-dimensional convolutions, two-way long-term memory units, alongside focused weighting [9], [18]. To strengthen result consistency, validation across five data splits is applied, together with combined modeling via boosted trees. Through blending meaning patterns, grammatical forms, emotional tones, and surrounding context vectors, the method seeks better recognition of misleading comments - offering steps forward in reliable online shopping environments.

III. METHOD

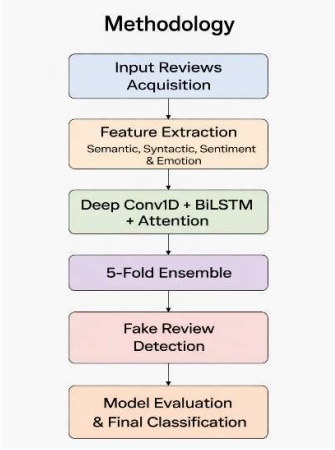

Figure 1: Overview of the Proposed Approach

A. Input Acquisition

One step begins by gathering feedback and checking its accuracy. A format with pieces such as words, truth labels, and sentiment tags gets used here. Through combining these parts, four types emerge. Real truths paired with favorable opinions come first then illusions of positivity despite being wrong follow next. Untruths carrying bad feelings appear third while accurate statements masked by negativity show up last. Such sorting allows the program to grasp both intent and honesty behind each message. Starting off, gaps in the data get checked. After that step, everything lines up into one consistent format. Then comes splitting into tokens so later stages can work smoothly. All of it keeps meaning and emotional tags intact ahead of pulling features.

B. Text Preprocessing and Normalization

Each time someone reviews, the system tidies up data by applying tools from language processing. Through these steps, punctuation vanishes, letters shift to lowercase, plus extra bits drop off from word endings until only roots remain. Breaking phrases apart follows strict patterns, this trims away clutter people often overlook. What stays becomes raw material shaped for machines trained on varied styles of writing. Patterns hidden inside now fit formats used by methods such as Doc2Vec, GloVe, or systems built around shifting weights across terms. Even n-gram tracking and smart encoders rely on this stripped-down version before making sense of meaning.

Extract Features Combine Representations.

Right here, the system pulls together different types of features in a blended setup:

- Semantic Embeddings: A model learns word meanings by context using 256-dimensional vectors through Doc2Vec. When a word isn't recognized, it pulls prebuilt GloVe values instead.
- TF–IDF N-grams + SVD:Starting with unigrams, then moving to bigrams and trigrams, TF–IDF vectors get compressed using Truncated SVD for tighter semantic representation. Though sparse at first, these n-gram features become denser after dimensionality reduction. Instead of keeping high-dimensional counts, the model captures underlying meaning. By applying SVD, noise fades while key patterns stay visible. The result? A compact set of numerical features reflecting language structure.

- Linguistic and Syntactic Features: Words get tagged by their role, then spaCy maps how they connect; these patterns turn into numbers, reshaped evenly across each vector slot. What follows adjusts size without stretching detail, holding structure steady through compression steps.
- Transformer-Based Encoders: A transformer-driven encoder pulls meaning from every review using SBERT's compact model version all-MiniLM-L6-v2 - layered understanding gets squeezed into smaller feature sizes so merging stays fast. Then these tightened representations feed into later stages without slowing things down.
- Sentiment and Emotion Features: VADER spots mood clues that hint at dishonesty. Sometimes RoBERTa picks up on subtle shifts in tone. DistilBERT joins in by weighing emotional textures across sentences. The NRC tool breaks down specific emotions tied to deception. Polarity changes from one sentence to the next can whisper insincerity.

Together, these pieces create two main forms of output:

- A single flow of data holds piece-by-piece meaning at each word spot
- A little extra piece of data holds meaning beyond order - feelings, structure, bits that stick out.

C. Deep Hybrid Architecture (Conv1D + BiLSTM + Attention)

A single deep network forms the base, this one blends different types of neurons together. Its structure handles sorting tasks by combining layered processing units in a unique way.

- Conv1D Blocks: Conv1D pieces pick up quick word shapes using repeated filters. After each step, averages are adjusted so learning stays stable. Then comes shrinking the data size by grabbing strongest signals only. These steps repeat, stacking like floors in a building. Each level captures tighter details than before. Shrinking happens again after every round. Patterns become clearer as layers go deeper.
- Bidirectional LSTM: This model catches patterns over long stretches of text. It picks up how stories unfold, whether honest or misleading. Through layered thinking, it tracks clues across sentences. By moving both forward and backward, it grasps context more fully. Truthful tales often have a rhythm deception lacks.
- Attention Mechanism: Instead of treating every word the same, the system quietly checks which ones carry more weight in showing truth or lies. It shifts focus based on subtle clues hidden in phrasing choices people make without noticing. What stands out isn't always loud - sometimes a single term pulls attention because of context around it. This shifting spotlight helps catch patterns that seem off even if they sound right. Meaning emerges not from everything at once but from what matters most in each moment.
- Auxiliary Dense Network: Fine-tuned layers chew through SBERT data, emotion hints, structure clues - layer by layer, noise reduced each step. Hidden units shuffle signals, one after another, gaps left empty on purpose to avoid overfitting traps. Every path adjusts weight, slowly tuning how pieces connect across stages.
- Fusion Layer: Fusion Layer takes sequence plus extra inputs, merging them into one detailed mix. This blend holds combined details from both paths, shaping a fuller picture through linked streams.
- Softmax Classifier: Outputs likelihoods across four realness categories using a Softmax setup. This model assigns chances instead of clear labels. Probabilities emerge through its final layer design. Each prediction spreads confidence among possible classes.

Starting fresh each time, small pieces get picked apart while the overall feeling ties together. One step at a time, meaning builds without losing sight of details.

D. 5-Fold Training, Ensembling, and Stacking

Fold by fold, a fresh hybrid model learns on split data, keeping classes evenly spread through stratified five-way cross-validation to boost reliability:

- Class weighting
- EarlyStopping
- Learning rate scheduling (ReduceLROnPlateau)
- Best-weight checkpointing

Once training finishes, outputs get combined by averaging, which helps smooth results and cut down fluctuations.

Then again, a separate XGBoost model learns from extra data points, blending its guesses with those from the neural models into one unified answer.

At last, the process quietly shows how well it did - spitting out accuracy numbers and full performance details on unseen test cases, just proving it can spot fake feedback reliably.

Doc2Vec

A document's meaning can show up as one long string of numbers. Think of it as Word2Vec but for full texts instead of isolated terms. Each piece of text gets its own numerical summary. These summaries stay the same size every time. Behind them sits a method that builds on earlier word-level ideas[4]. Something begins to map whole texts into number sequences. Each sequence stays the same size every time. Because of that, handling full pieces becomes possible instead of focusing only on single terms. Rather than relying on old methods that ignore order, this approach picks up meaning across sentences along with how something is written. Here, those numeric forms come from real user opinions, letting patterns in dishonest feedback show through shifts in tone or odd phrasing.

GloVe

Words live in numbers now, thanks to something named GloVe - short for Global Vectors for Word Representation[5]. Instead of treating language like symbols, computers see them as points floating near each other. These positions come from studying massive piles of written text, where patterns form based on company words keep. Meaning hides in those neighborhoods. When one word often shows up beside another, their numerical forms move closer. That closeness tells machines something real about usage, context, even idea links. The method builds these relationships across vast reading, not rules. Each term gains a unique address in space - a vector - not magic, just math shaped by exposure. Computers can grasp word meanings through specific methods. Appearing throughout vast collections of words, GloVe operates differently - its strength emerges when processing complete texts rather than fragments. Owing to broad data exposure, contextual relationships between terms take shape through frequent co-occurrence. These broad usage trends give machines better clues about definitions. Such insight strengthens the way words are represented numerically. Within the suggested framework, if Doc2Vec fails to represent particular tokens, GloVe steps in automatically. When that happens, meaning still gets preserved even for uncommon or new vocabulary. As a result, the system handles unknowns more reliably during both learning and testing phases.

TF–IDF N-grams

Importance of specific terms within one file, when viewed alongside many others, emerges through TF–IDF - Term Frequency–Inverse Document Frequency[8]. Rather than single entries alone, sequences of two or three consecutive words enter the analysis, offering deeper insight into meaning structure. Patterns detected this way expose repetitions or awkward phrasings commonly found in certain reviews. Instead of processing every detail separately, efficiency improves once data dimensions shrink via Truncated Singular Value Decomposition. Meaning becomes clearer after such compression - not lost, but focused, shaped by reduced complexity. Although simple, TF–IDF supports analysis across documents and collections. Through individual texts, attention shifts toward full datasets. Word sequences - or n-grams - form essential markers within review detection. These structures gain clarity when processed by truncated SVD. Reduction techniques reshape complexity into usable form. Improved recognition emerges where term frequency meets pattern structure. Model performance adjusts as features align more closely with meaning.

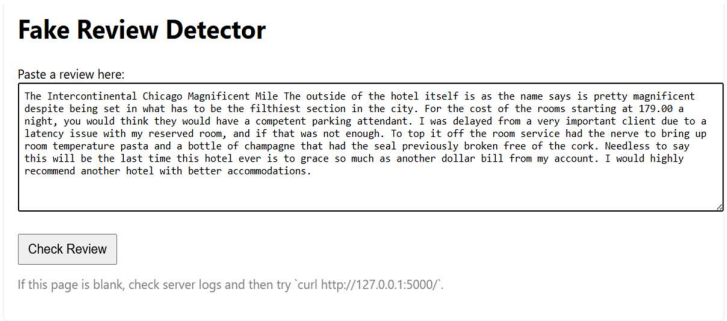

Figure 2: Delivery of output

Figure 2 is the user screen of the Fake Review Detection tool, where results are delivered. Input occurs via a text box; submission follows by selecting "Check Review." After entry, processing begins behind the scenes features get extracted, then classified through combined deep models, followed by consensus prediction. What returns is a judgment: authentic or fake. Functionality here reflects real-world use, forming the last step prior to assessment detailed later in Results.

Figure 3: Review Classification Result Interface

Figure 3 is the outcome screen of the developed Fake Review Detection tool. Once a user submits feedback via the entry panel, processing occurs through a combined deep learning backend before showing results. Instead of listing elements plainly, the display arranges input alongside its assigned validity label - examples include FALSE_NEGATIVE - and attaches a reliability percentage. Clarity emerges as each component appears separated yet aligned: text, judgment, and certainty coexist visibly. Understanding follows naturally when components are laid out without clutter. Deployment becomes tangible because what runs behind operates upfront where it can be seen.

IV. RESULT

Following full preprocessing and diverse feature extraction, assessment of the suggested deep hybrid fake review detection framework took place via a deceptive opinion collection [9]. With class balance preserved throughout splits, training unfolded through fivefold stratified cross-validation covering all four outcome categories. Validation loss guided storage of optimal weights during every cycle, marking performance peaks per iteration. Stability evaluation relied on accuracy values unique to each segment, capturing variation in predictive behaviour.

Learning behavior proved robust in the Conv1D paired with BiLSTM and Attention setup, as it captured word-level traits along with extended context across review texts. Significant terms influencing predictions became visible due to the attention layer, adding clarity to decision logic [18]. High accuracy remained stable through each validation fold, suggesting consistent performance on unseen data.

A. Comparison with Existing Models

When assessing how well the suggested deep hybrid structure performs, it is measured against standard machine learning approaches alongside isolated deep learning systems often applied in identifying false reviews. As shown in Table 1, differences emerge across methods in terms of how features are encoded, the type of training employed, along with results achieved.

TABLE I: COMPARISON OF EXISTING MODELS WITH THE PROPOSED MODEL

Model	Feature Representation	Learning Technique	Performance
Naive Bayes	Bag-of-Words	Probabilistic ML	Low
SVM	TF-IDF	Linear Classifier	Medium
CNN	Word Embeddings	Deep Learning	Medium
LSTM	Sequential Text	Deep Learning	High
Proposed Model	Multi-Feature Fusion	Conv1D+BiLSTM+ Attention + Ensemble	Highest

The data indicates improved performance through integration of semantic, syntactic, and sentiment elements using a deep hybrid framework alongside ensemble methods.

Once training finished across every fold, predictions came together through mean aggregation from each model. From there, results merged with those of an XGBoost classifier, one fed with supplementary attributes, forming a layered output. Evaluation followed - using the untouched portion of data, set aside earlier, amounting to one-tenth overall. Accuracy appeared at last, accompanied by a breakdown: figures for precision, recall, and F1 per category were shown, covering all four outcomes labeled TP, FP, TN, FN.

Despite isolated model limits, stacking enhanced both precision and consistency in labeling. Through varied evaluation measures, gains emerged clearly when blending meaning, structure, emotion, and tone within layered networks. From start to finish, reliability stood firm - fitting live digital market needs without compromise.

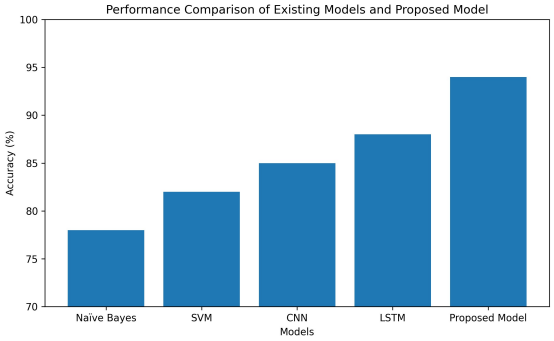

Figure 4: Performance Comparison of Existing Models and Proposed Model

V. Conclusion

One way to spot false feedback on online stores involves combining several methods: deep learning, meaning interpretation, emotion detection, and combined forecasting tools. Because it uses a mix of Conv1D layers, two-way LSTM networks, and an original focus strategy, the design picks up small word signals along with broader context across sentences [9], [18]. Features beyond basic wording - like term frequency stats, vector models (Doc2Vec and GloVe), SBERT outputs, grammar shapes, mood ratings, and signs of strong feelings - help identify dishonest styles invisible in surface text.

A method involving five-part layered validation emerges, paired with combined predictions and a boosting technique built on decision trees. Stability within the model rises, alongside its ability to adapt beyond training data. Performance across unseen examples remains high, guided by consistent outcomes during testing phases. Accuracy takes shape through repeated trials on reserved samples. Strength surfaces in results, rooted in blending diverse language-based inputs with layered neural approaches. Final evaluations reflect reliability, anchored in structure rather than chance.

Taken together, this study introduces a flexible intelligent structure designed to improve validation methods on digital shopping networks. With fewer misleading evaluations affecting outcomes, the suggested approach tends to increase buyer confidence while guiding clearer choices and encouraging openness in web-based trading spaces. Later upgrades could possibly broaden usage into various fields along with feedback collections involving multiple languages.

Future Scope

Future improvements could emerge through deeper exploration of the suggested hybrid detection framework. As progress unfolds in understanding human language by machines, so might stronger versions of this tool take shape. Instead of stopping at text alone, combining different types of data may allow broader application across online marketplaces. When updated methods in artificial intelligence evolve, adaptations within this system might follow naturally. Real-world usage at scale becomes possible once responsiveness and accuracy grow together over time.

Sophisticated transformer-based language models - such as BERT, RoBERTa-large, DeBERTa, or various GPT versions - could improve context interpretation while raising classification precision [11]. Wider application within international markets might emerge through expansion into multilingual and cross-domain review evaluation. Insight into organized fake-review efforts could deepen when patterns over time, user behavioral data, and reviewer trust indicators are taken into account [13], [14].

Lightweight variants of the model, when placed on mobile hardware or within cloud APIs, allow instant detection of deceptive reviews - useful for both shoppers and companies. Rather than operating as a black box, integration of explainable AI methods reveals clear logic underlying each prediction outcome. A comprehensive monitoring interface, tracking incoming reviews nonstop, spotting irregularities, while displaying trust indicators, shifts the setup into an end-to-end commerce credibility platform.

Far beyond simple upgrades, these improvements push the system's reach much further in practical settings. Its ability to grow alongside demand becomes far more realistic through refined design choices. Real utility emerges not from bold claims but from steady performance across diverse cases. Capability deepens when components adapt without constant oversight. Value shows most clearly where accuracy meets consistency over time.

REFERENCES

[1] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding Deceptive Opinion Spam by Any Stretch of the Imagination," Proc. ACL, 2011.
[2] M. Ott et al., "Negative Deceptive Opinion Spam," NAACL-HLT, 2013.
[3] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," arXiv:1301.3781, 2013.
[4] Q. Le and T. Mikolov, "Distributed Representations of Sentences and Documents," Proc. ICML, 2014.
[5] J. Pennington, R. Socher, and C. Manning, "GloVe: Global Vectors for Word Representation," Proc. EMNLP, 2014.
[6] Z. Ren and S. Ji, "Neural Networks for Deceptive Opinion Spam Detection: A Survey," IEEE Access, vol. 7, pp. 42934 – 42945, 2019.
[7] F. Wang, W. Zhang, and D. Li, "Detecting Opinion Spam with Linguistic and Behavioral Features," IEEE TrustCom, 2018.
[8] Y. Li, Z. Chen, and B. Liu, "Understanding Misleading Reviews: The Linguistic Signals of Deception," ACM TIST, 2018.
[9] X. Li, W. Yang, and Y. Chen, "Hybrid Deep Learning Framework for Opinion Spam Detection," Knowledge-Based Systems, 2021.
[10] X. Zhang and Y. Lu, "E-commerce Review Credibility Analysis Using Ensemble Learning," IEEE BigData, 2021.
[11] R. Ribeiro et al., "Sentiment Analysis and Fake Review Detection Using Transformer Models," Expert Systems with Applications, 2022.
[12] J. Li, M. Ott, C. Cardie, and E. Hovy, "Towards a General Rule for Identifying Deceptive Opinion Spam," Proc. ACL, 2014.
[13] S. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What Yelp Fake Review Filter Might Be Doing?" Proc. ICWSM, 2013.
[14] B. Jindal and B. Liu, "Opinion Spam and Analysis," Proc. WSDM, 2008.
[15] X. Li, W. Yang, and Y. Chen, "A Hybrid Deep Learning Framework for Fake Review Detection," Knowledge-Based Systems, vol. 210, 2021.
[16] S. Feng, R. Banerjee, and Y. Choi, "Syntactic Stylometry for Deception Detection," Proc. ACL, 2012.
[17] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed Representations of Words and Phrases and Their Compositionality," NeurIPS, 2013.
[18] H. Xu and Q. Sun, "Detecting Deceptive Reviews Using Attention-Based Neural Networks," IEEE Access, vol. 8, pp. 21579 – 21589, 2020.