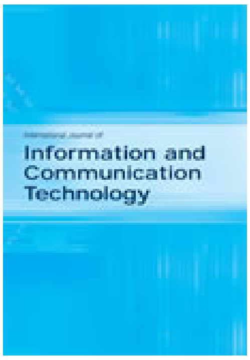


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Transformers-based feedback analysis of e-commerce: a focused study on quality assessment of agriculture products

Wenrui Xu

DOI: 10.1504/IJICT.2025.10071015

Article History:

Received:	16 March 2025
Last revised:	31 March 2025
Accepted:	31 March 2025
Published online:	27 May 2025

Transformers-based feedback analysis of e-commerce: a focused study on quality assessment of agriculture products

Wenrui Xu

Electronic Commerce,
Guangdong Polytechnic of Science and
Technology School of Business,
Zhuhai, Guangdong, 519090, China
Email: xuwrr001@163.com

Abstract: Ensuring the quality of agricultural products in e-commerce is a significant challenge due to product variability and the absence of direct inspection before purchase. Customer reviews serve as a critical source of information, offering insights into product freshness, packaging, and overall satisfaction. This research focuses on the agricultural product domain, where quality plays a pivotal role in ensuring consumer trust and satisfaction. Harnessing the power of large language models (LLM), this study investigates the application of state-of-the-art transformer-based models for analysing customer feedback. The research utilises fine-tuned BERT and RoBERTa models to classify and predict product quality based on sentiment and contextual analysis of user reviews. The findings highlight the remarkable performance of these models, with RoBERTa achieving the highest accuracy of 99%. This study signifies the growing importance of AI and LLMs in enhancing e-commerce practices, particularly in domains like agriculture, where product quality assessment is paramount.

Keywords: artificial intelligence; e-commerce; agriculture; transformers; deep learning; sentiment analysis; natural language processing; NLP; machine learning.

Reference to this paper should be made as follows: Xu, W. (2025) 'Transformers-based feedback analysis of e-commerce: a focused study on quality assessment of agriculture products', *Int. J. Information and Communication Technology*, Vol. 26, No. 14, pp.62–86.

Biographical notes: Wenrui Xu is a researcher at Guangdong Polytechnic of Science and Technology, and his fields of interest in electronic commerce, digital marketing, and business intelligence. In e-commerce, he does research on how to structure an e-commerce strategy, how to analyse consumer behaviour, and how to leverage big data to aid in business decision making.

1 Introduction

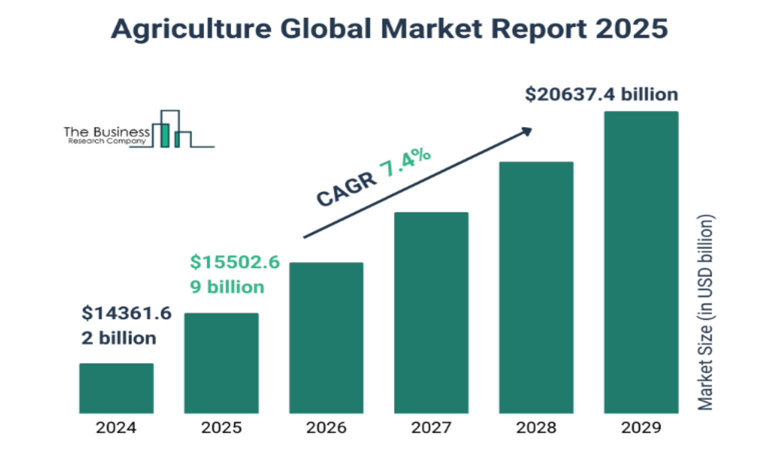
Agriculture is a cornerstone of the global economy, playing a vital role in the economic security of nations by serving as a primary source of food and raw materials. It significantly contributes to job creation and income, especially in rural areas where it supports the livelihoods of a substantial portion of the population. As the world's population continues to grow, the agricultural sector's role in ensuring food security and promoting economic stability becomes increasingly critical, shaping the development agendas of nations worldwide (Lin, 2020).

Globally, agriculture employs over 9.9 billion people, accounting for approximately 28% of the world's workforce. This proportion varies significantly between developed and developing nations. For instance, in many developing countries, particularly in regions like Sub-Saharan Africa, more than 60% of the population is engaged in agriculture, contributing about 23% to the region's GDP. The economic impact of agriculture is particularly substantial in low and middle-income countries, where it can account for 25% to 30% of GDP. Beyond direct employment, the sector also bolsters industries such as manufacturing and services by producing agricultural goods and fostering related businesses.

However, maintaining the quality of agricultural products presents substantial challenges. Traditional methods, such as physical inspections and expert evaluations, are effective but labour-intensive and often struggle to keep pace with the scale and rapid development of modern agricultural operations. In response, the integration of AI offers a groundbreaking solution. Advancements in AI, particularly through natural language processing (NLP) and machine learning (ML), enhance the efficiency and accuracy of quality assessments (Khrais, 2020). AI technologies enable the rapid analysis of vast amounts of data – from environmental conditions to consumer feedback – providing a detailed and nuanced understanding of product quality that surpasses traditional methods. This evolution not only aids in better decision-making but also promotes a more sustainable and responsive agricultural framework, essential for the ongoing economic stability and growth of nations (Talha et al., 2023).

Agriculture plays a pivotal role in shaping the economies and societies of nations worldwide. It serves as the backbone of many economies, providing sustenance and livelihoods to billions. The agricultural sector not only ensures food security but also contributes significantly to the export and import trade, strengthening global economic ties. In the modern era, the evolution of agricultural practices and technologies revolutionised how crops are cultivated, harvested, and distributed, driving significant economic and societal transformations (Vanneschi et al., 2018). The sale of agriculture-based products has expanded from traditional markets to modern e-commerce platforms. This transition to online marketplaces has provided farmers and producers with an opportunity to reach broader audiences and improve profit margins (Alzahrani et al., 2022). Consumers now have access to a diverse range of agricultural products at their fingertips, from fresh produce to processed goods. However, with the rise of online shopping, evaluating the quality of products remotely has become a significant challenge for consumers and sellers alike (Daza et al., 2024). Advancements in artificial intelligence (AI) have paved the way for innovative solutions in addressing this challenge. AI-driven tools enable efficient and reliable analysis of consumer-generated data, such as reviews and feedback, to assess product quality (Naz et al., 2024). Sentiment analysis, a prominent AI technique, focuses on extracting subjective opinions and emotions from textual data. By analysing customer reviews, AI systems can detect the sentiment behind user opinions, providing valuable insights into product satisfaction, quality, and usability (Nawaz et al., 2021). AI has emerged as a transformative force in analysing user-generated content from e-commerce platforms, particularly for agriculture-based products. In the e-commerce landscape, a user's product selection often reflects their personality traits and lifestyle preferences, offering valuable insights into their individual needs and behaviors (Naz et al., 2024). User reviews serve as a rich source of information, reflecting the experiences and satisfaction levels of consumers with agricultural goods such as fresh produce, seeds, fertilisers, and equipment. By applying AI techniques to this data, businesses can uncover patterns and trends that are not immediately apparent (Ahmad et al., 2023). Advanced models excel at processing the complex language of user reviews, identifying nuances such as tone, sentiment, and context (Huang et al., 2023). For agricultural products, this analysis can reveal critical insights, such as issues related to freshness, packaging, delivery, or product authenticity, thereby empowering sellers to make data-driven improvements and consumers to make informed purchasing decisions (Benos et al., 2021).

Figure 1 Analysis of market trend of agriculture based products growth (see online version for colours)



The practice of purchasing goods through the internet has been on the rise in the last decade, buoyed by more effective connectivity, ease in making transactions and a wide range of products available on the internet. Current trends in e-commerce have consumers using outlets prominently in e-business including amazon, eBay, Alibaba, Flipkart and others such as Lazada (Cravero et al., 2022). They enable ease of use, secure payment methods and efficient delivery Channel hence becoming the preferred mode of shopping. The notable rise in the market size year over year suggests significant investments and advancements in agricultural technologies and practices. This trend points towards an increasing demand for agricultural products and potentially improved productivity in the sector, driven by innovations and an expanding global population. The graph underscores the critical role of agriculture in the global economy and the dynamic changes anticipated in the coming years, also defined in Figure 1. Furthermore, social commerce such as Instagram Shopping and Facebook Marketplace have emerged developing new ways of user engagement with products by providing social shopping experience in an integrated social media marketplace. The analysis of these sites shows that their consumers are driven by factors like simplicity, availability and variety of choices, reliability, and post purchase services (Noor and Islam, 2019).

1.1 Research contributions

In this study, we explore the application of advanced AI methods to improve the quality assessment of agriculture-based products sold on e-commerce platforms. Specifically, we leverage the masked language model RoBERTa to classify users' reviews by extracting meaningful features from a domain-specific dataset. This research aims to enhance the reliability and efficiency of product quality evaluations, contributing to improved consumer experiences and informed purchasing decisions in the agricultural e-commerce domain.

1.2 Paper organisation

The rest of the paper follows as: Section 2 presents the comprehensive analysis of existing studies based on traditional approaches to advanced transformer-based methods. Section 3 presents the problem statement along with formulation. Section 4 highlights the proposed methodology to conduct this study, sharing experimental setup. Section 5 provides a comprehensive analysis of results along with descriptive and predictive outcome. Section 6 shares the conclusion of our findings with proposing future work.

2 Related work

The review analysis of agricultural products by users is significant research that focuses on enhancing consumers' opinions of different crops in the farming sector. To improve this analysis's effectiveness, the application of modern AI technologies like ML, ensemble learning, deep learning (DL), and transformer models have been investigated, as taxonomy defined in Figure 2. Furthermore, Table 1 shares the summary of existing studies.

Figure 2 Applied taxonomy of existing studies (see online version for colours)

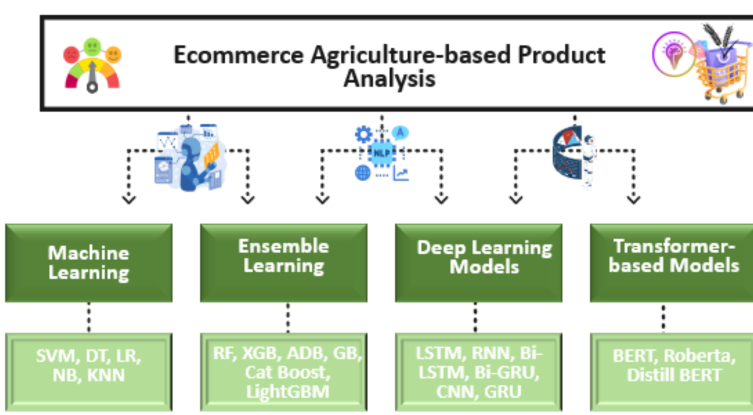


Table 1 Existing studies summary

Ref	*Year*	*Model*	*Dataset*	*Features*	*Results (%)*
Noor and Islam	2019	Naïve Bayes, J48	Women's e-commerce reviews	TF-IDF	89
Yang et al.	2020	CNN, BiGRU	Book reviews	Count vectorisation	92
Singh and Sarraf	2020	RF	User reviews	Textual features	88
He et al.	2020	BiLSTM, CNN	Chinese reviews of agricultural products	FastText, and Word2Vec	89
Xu et al.	2020	NB	Review sentiment	Textual features	70
Benos et al.	2021	ANN,SVM	Crop dataset	Textual features	78
Fidyanti et al.	2021	SVM	Product selection data	POS tagging	89
Rong et al.	2021	CNN	Product review	Word2Vec	88
Cravero et al.	2022	SVM, DT, NB	Historical records of soil	Textual features	80
Arobi et al.	2022	KNN, RF, LR	E-commerce customer review	Textual feature	87
Jahan et al.	2022	NB, KNN	Consumer comments on Daraz	Ngram, bigram, unigram	90
Alghazzawi et al.	2023	ERF-XGB	IMDB datasets	Harris hawk optimisation	92
Cao et al.	2022	SVM, NB	Customer feedback	Textual features	90
Liu	2024	RNN, BiGRU, and Bert-BiLSTM	Product reviews	Word embeddings	95
Shang et al.	2014	CNN, BERT, BiLSTM	Reviews of different products	FastText, and Word2Vec.	94
Ali et al.	2024	CNN, BERT, XLNet	Amazon reviews	TF-IDF, BOW	89
Xiang et al.	2024	BERT	Reviews of different products	Word embeddings	90
Taneja et al.	2024	BERT	Women clothing ecommerce dataset	Word embedding	95
Alsaedi et al.	2024	CNN, RNN, LSTM	Customer reviews	Word embedding	93
Bellar et al.	2024	BERT, LSTM, CNN	Reviews of different products	FastText, and Word2Vec	91

2.1 Existing studies of ML

In the review analysis, the agriculture products users' review analysis concentrates on utilising higher technologies like ML and NLP to enhance decision-making regarding product development. According to the importance of ML for agriculture, the application of these technologies in crops, water soil, and animals was discussed. Still, it paid relatively little attention to practical concerns with the adoption of these technologies alongside the corresponding ethical considerations related to utilising them in the agricultural context (Benos et al., 2021). Some of the problems that the analysis also highlighted include data inaccessibility, and poor rural bandwidth when it comes to big data in agriculture ML. Though these issues were discussed, the considerations did not offer any solution or examples of the applications of ML (Cravero et al., 2022).

The review in this paper looked at ML technology applications in agricultural supply chains, especially about sustainability and decision-making. Nevertheless, it did not include the empirical testing of the proposed frameworks and omitted several tangible challenges of applying ML in diverse agriculture contexts, which are necessary for evaluating the possibilities of its implementation (Noor and Islam, 2019). Further, a new sentiment analysis model, SLCABG, is presented for Chinese e-commerce reviews using sentiment lexicons and DL. While encouraging this model did not enter comparative analysis with more sophisticated models and did not consider the problems related to the use of informal language, which is critical when assessing user feedback in agriculture (Yang et al., 2020). Finally, the application of ML and text mining to stock market prediction, particularly news data was also addressed. While noting that the fragment has been used more and more in predictions, it said that current more sophisticated NLP techniques like transformers have not been applied fully yet. This also meant that there was no comparison between text transformation techniques and feature representation techniques which are crucial when analysing agricultural product reviews (Cao et al., 2022).

2.2 Existing studies of ensemble learning

Of all the tasks related to consumer feedback analysis, user review analysis of agricultural products is paramount important. The state-of-the-art ML education focused on several views in Agriculture 4.0 and provided essential information about the methodologies of review analysis. The novelty was an ensemble ML technique that was proposed to improve network security in the context of Agriculture 4.0. Although it was promising, the method did not include the evaluation of various datasets and did not meet the problem of computational complexity regarding ensemble approaches (Singh and Sarraf, 2020). The problem of agricultural futures price forecasting was addressed in one work by using the ensemble empirical mode decomposition (EEMD) combined with long short-term memory (LSTM) networks. This hybrid model enhanced short-run forecast accuracy but did not investigate the effects of decomposition levels and had a sound long-run validation (Alghazzawi et al., 2023). An integrated system of SVM and ERNIE was built to forecast multi-layer analysis in an agriculture-leading basin in China. While it yielded enhanced differentiation compared with conventional methodologies, its feasibility under different weather conditions and applicability to various locations was unknown (Fidyanti et al., 2021). They also demonstrated that a DL-based approach for mining product innovation ideas from online reviews generates more ideas than the traditional approach. Still, it failed to consider the effect of fine-grained domain differences in reviews and did not test it at a larger scale on different goods categories (Arobi et al., 2022). Likewise, the GBSVM model that combined results from gradient boosting and support vector machines was introduced for the classification of structured sentiment of unstructured reviews. The gain was big, but the model lacked one parameter – computational complexity and one aspect – difference in different languages (Jahan et al., 2022). Such developments proved that, in the sphere of agriculture, ML could be successfully implemented. However, the extensions of the evaluation to multiple datasets, considering the computational issues, and the scalability were crucial to guarantee the stability and the generalisation of the findings, especially in the study of the reviews of agricultural products.

2.3 Existing studies of DL

Agriculture products users' review analysis entailed sophisticated techniques such as DL to analyse and understand users' feedback to enhance production decision-making in the agricultural sector. This model was suggested for sentiment analysis for e-commerce product reviews and yielded an accuracy of 95.5%. (Liu, 2024). Prior studies focused on DL applications failed to consider issues related to data quality that has a direct poor correlation with the model. It also did not cover functional issues in trying to apply such a scale of concepts in agriculture (Shang et al., 2014). Previous studies on DL in cattle farming did not consider how to scale up increased production and did not find any concern with environmental conditions of light and sound that help the model's accuracy (Rong et al., 2021). There were several proposed approaches in applying DL to smart fish farming and their respective drawbacks: Few of them paid attention to real-world implementation issues involving DL, and there was an absence of discussions on the constraints in data diversification (He et al., 2020). Apart from integrating online lexicons with DL for Chinese e-commerce reviews, sentiment analysis lost comparative study with other approaches and did not address problems associated with informal expressions in user feedback (Xu et al., 2020).

2.4 Existing studies of transformer-based learning

We investigated methods of analysing user feedback by using extended quantitative methods. The statistical, ML and DL techniques were evaluated on six models using agrometeorological data collected from online platforms. The TCNN model provided the best general performance but failed to compare its performance across various climatic conditions and did not explore the overfitting of DL models (Ali et al., 2024). Another study explained the integration approaches for predicting the binary sentiment of the user's reviews in the Brazilian Portuguese language. Proper TLMs yielded better performance than the conventional methods including bag-of-words and other deep models such as CNNs and LSTMs. However, it failed to consider the computational cost in terms of the computations required and the problems associated with applying these models to other NLP tasks or other languages (Xiang et al., 2024). There was a technique that used rating scores as an alternative to assessing product popularity. As this was a new approach it did not consider the feature variability of different categories of products and did not have an approach for the large dataset cases (Taneja et al., 2024). Additionally, the focus was placed on the efficiency of NLP tasks using transformer-based pre-trained models like BERT and GPT.

However, absolute and relative performances across languages were missing, and the resource demand for training and deployment was not discussed (Alsaedi et al., 2024). One work integrated dictionary-based, rule-based, and transformer-based methodologies for agricultural entity identification with high precision, recall, and F1-score. Nevertheless, the results posed issues concerning framework scalability in parallel to concerns about variability across the sub-domain of agriculture (Bellar et al., 2024).

These results indicated the possibility of the superior performance of sophisticated techniques for processing user feedback and underscored the necessity for future research to address the issues of increasing problem size, reduced computation expenses, and questions specific to specific fields.

3 Problem statement and formulation

Let $R = \{r_1, r_2, \ldots, r_n\}$ denote a set of N user reviews for an agricultural product on an e-commerce platform, where each review $r_i = (w_1, w_1, \ldots, w_{T_i})$ is a sequence of T_i tokens. The quality of product Q, is a target variable that can be modelled as a discrete categorical variable $Q \in \{q_1, q_2, \ldots, q_K\}$. The objective is to learn a mapping $f: R \to Q$ such that $Q = f(R)$, where the objective f is a transformer-based DL model. The model first encodes each review r_i into a contextualised embedding $E_i = TransformerEncoder(r_i)$ where $E_i \in \mathbb{R}^{d*T_i}$ and d is the embedding dimension. These embeddings are aggregated into a fixed-dimensional function $H = g(E_1, E_2, \ldots, E_n)$, where $g: \mathbb{R}^{d*\sum_{t=1}^{N} T_i} \to \mathbb{R}^d$ is an aggregation function for attention mechanism. Finally, H is passed through a classifier as: $\mathbb{R}^d \to Q$ to predict Q.

The problem is constrained by challenging as noise in R, where reviews may contain irrelevant tokens, slang or misspellings, leading to suboptimal embeddings E_i, contextual understanding, requiring the model to capture semantic relationships $TransformerEncoder(r_i)$ that distinguish between phrases, furthermore, dealing imbalanced data, where the distribution of Qskewed, necessitating techniques like weighted loss functions $\mathcal{L}(Q, \dot{Q})$ to handle underrepresented classes and scalability as the model must effectively process large-scale reviews sets R in real-time. The goal is to optimise f such that it minimises the prediction error $\|Q - \dot{Q}\|$, where $\dot{Q}$ is the predicted quality, while addressing these constraints to enable robust and scalable quality assessment of agricultural products in e-commerce platforms.

4 Research proposed methodology

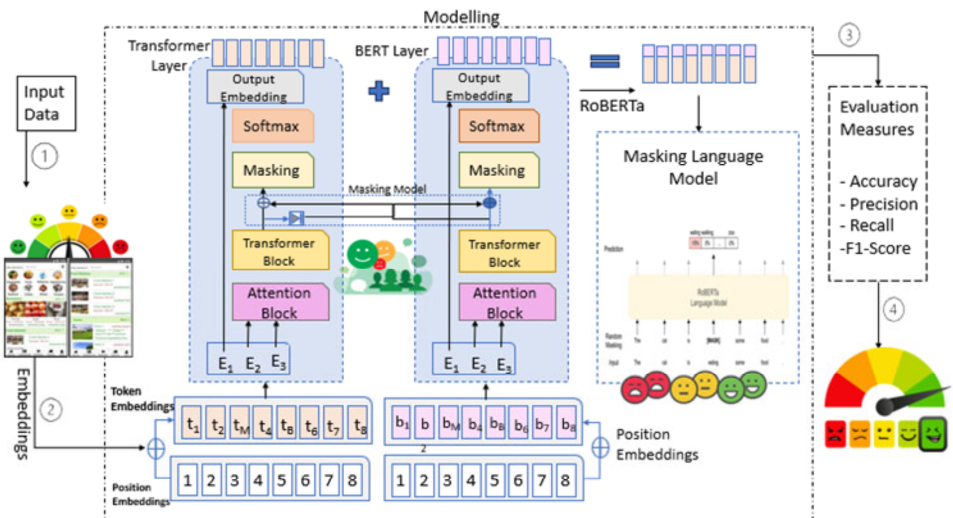
In this section, we present the strategic steps to fulfil our research aims. These steps, defined in Figure 3, involve data cleaning and sorting before choosing useful features that lead to building ML models and testing their accuracy. Together these methods create reliable results. The specific features of our problem domain influence every step of our methodology which results in an efficient and scalable solution.

4.1 Data preprocessing

Let $R = \{r_1, r_2, \ldots, r_n\}$ represent the raw set of N user reviews, where each review r_i is a sequence of characters. The preprocessing pipeline transforms R into a clean and structured dataset $\dot{R} = \{\dot{r}_1, \dot{r}_2, \ldots, \dot{r}_n\}$. First, missing values in R are handled by removing them, ensuring R is complete. Next each review r_i undergoes text cleaning: removal of stop words S, punctuation P, and non-alphabetic character, resulting in

$\dot{r}_i = clean(r_i, S, P)$. Lemmatisation is the applied to reduce words to their base form (Naqvi et al., 2023), $lemmatizer(\dot{r}_i)$ followed by tokenisation, which splits each review into a sequence of tokens $\dot{r}_i = (w_1, w_2, ..., w_{T_i})$, where w_j is the j^{th} token and T_i is the tokenised length of the i^{th} review. The pre-processed dataset $\dot{R}$ is thus represented as $\dot{R} = \{tokenize(lemmatize(clean(r_i, S, P))) \mid i = 1, 2, ..., N\}$.

Figure 3 Proposed research proposed methodology (see online version for colours)



4.2 *Proposed model*

RoBERTa-large is an advanced variant of the RoBERTa model, featuring significantly larger architecture with more parameters, enabling it to capture even richer contextual representations of text. It is designed to handle complex natural language understanding tasks, such as quality assessment of agricultural products based on user reviews. The model leverages a deep Transformer encoder with enhanced feature extraction capabilities, making it highly effective for downstream tasks (Ahmed et al., 2023).

4.3 *Feature extraction*

4.3.1 *Input representations*

The input to RoBERTa$_{\text{[LARGE]}}$ is a sequence of tokens $\dot{r}_i = (w_1, w_2, ..., w_{T_i})$ derived from pre-processed user reviews. Each token w_j is converted into a dense vector representation using a combination of token embeddings, positional embeddings, and segment embeddings. The input embeddings for the j^{th} token is computed as computed using equation (1):

$$x_j = W_t.one - hot(w_j) + W_p.pos(j) + W_s.seg(j) \tag{1}$$

where $W_t \in \mathbb{R}^{d*|V|}$ represents token embeddings matrix, $|V|$ is vocabulary size, d is embedding dimension, $W_p \in \mathbb{R}^{d*T_{\max}}$, $W_s \in \mathbb{R}^{d*2}$ shows positional and segment embedding matrix. $pos(j)$ represent positional encoding for the j^{th} token, which provides information about the token's position in the sequence. $seg(j)$ shows the Segment ID, used to distinguish between multiple segments.

The resulting input embeddings $X = [x_1, x_2, ..., x_{T_i}]$ are passed through the transformer encoder to generate contextualised representations.

4.3.2 Masked language modelling

RoBERTa$_{\text{[LARGE]}}$ is a pre-trained using a masked language modelling objective, where a subset of tokens $M \subset \{1, 2, \ldots, T_i\}$ is randomly masked. The model is trained to predict the masked token w_m where $m \in M$ based on the context provided by the unmasked tokens. The probability distribution for the masked token w_m is computed using equation (2):

$$P\left(w_m \mid \dot{r}_i \setminus w_m\right) = \text{softmax}\left(E_i^{(m)} \cdot W_t^T\right) \tag{2}$$

where $T_{\max}$ shows maximum sequence length, $E_i^{(m)} \in \mathbb{R}^d$. Contextualised embedding for the masked position m, generated by the transformer encoder. W_t^T token embedding matrix, shared with the input embedding layer.

The masked language modelling (MLM) objective boost the model to learn deep bidirectional contextual representations, as it rely on both left and right context to predict the masked tokens.

4.4 Transformer encoder architecture

4.4.1 Multi-head self-attention

The core component of the transformer encoder is the multi-head self-attention (MHSA) mechanism, which enables the model to capture dependencies between tokens in the sequence. The input embeddings $X = [x_1, x_2, ..., x_{T_i}]$ are passed through H parallel attention heads, For each head h, the query Q_h, key K_h, and value V_h metrices are computed as in equation (3):

$$Q_h = X.W_h^Q, \quad K_h = X.W_h^k, \quad V_h = X.W_h^V \tag{3}$$

where $W_h^Q, W_h^k, W_h^V \in \mathbb{R}^{d*d_k}$ are learnable weights.

The attention output for the head h is computed using equation (4):

$$head_h = \text{Softmax}\left(\frac{Q_h K_h^T}{\sqrt{d_k}}\right) V_h \tag{4}$$

where $d_k = d/H$ is dimension of each head.

The outputs of all heads are concatenated and projected, as in equation (5):

$$MHSA(X) = Concat\left(head_1, ..., head_H\right).W^{\circ} \tag{5}$$

where $W^{\circ} \in \mathbb{R}^{d*d}$ learnable weight matrix of MHSA, $W_1 \in \mathbb{R}^{d*d_{ff}}$, $W_2 \in \mathbb{R}^{d_{ff}*d}$ intermediate weight dimensions

The MHSA mechanism allows the model to attend to different parts of the sequence simultaneously, capturing both local and global dependencies.

4.4.2 Feed-forward neural network

The output of the MHSA is passed through a position wise feed-forward neural network (FFN) which consists of two linear transformations with a GELU activation function.

4.4.3 Layer normalisation and residual connections

Each sub-layer (MHSA and FFN) is followed by layer normalisation and a residual connection. Layer normalisation stabilises and accelerates training by normalising the activations of a layer across the feature dimension. This ensures that the activations have a consistent mean and variance, reducing internal covariate shift and improving convergence. For each layer as input $X \in \mathbb{R}^{T_i*d}$ defined as in (6)

$$LayerNormK* = (X) = \gamma \odot \frac{X - \gamma}{\sqrt{\sigma^2 + \in}} + \beta. \tag{6}$$

where $\gamma = \frac{1}{d}\sum_{j=1}^{d} X_{ij}$ is mean along the feature dimension, $\sigma^2 = \frac{1}{d}\sum_{j=1}^{d}\left(X_{ij} - \gamma\right)^2$ is variance along the feature dimension. $\gamma, \beta \in \mathbb{R}^d$ is learnable parameters. $\in$, and X are small constant for numerical stability, with input to the sub-layer

Residual connections address the vanishing gradient problem by adding the input of a layer directly to its output. By combining the input and output, residual connections enable the training of very deep networks without degradation in performance. For a sub-layer defined in equation (7).

$$X_{out} = LayerNorm(X + SubLayer(X)) \tag{7}$$

where *SubLayer* represents either MHSA or FFN, *SubLayer*(X) is the output of sublayer. X + *SubLayer*(X) is residual connection. These components stabilise training and improve gradient flow.

4.4.4 Output representation

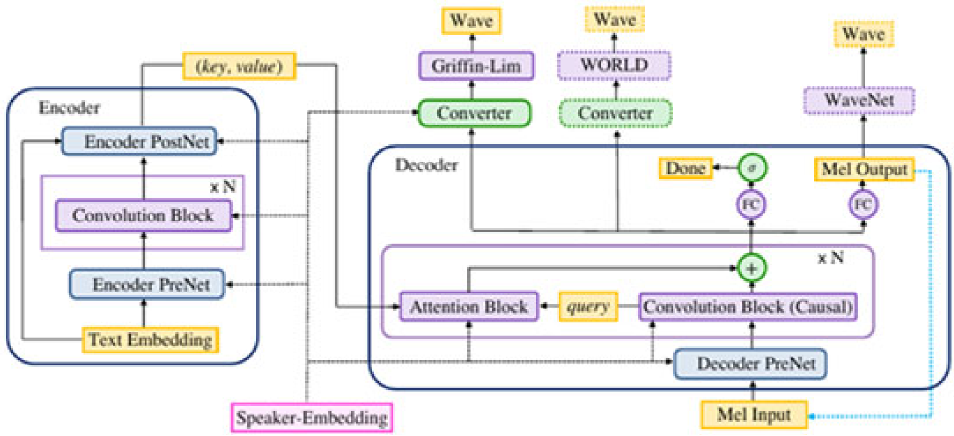
The final output of the transformer encoder is a sequence of contextualised embeddings $E_i = [E_i^{(1)}, E_i^{(2)}, ..., E_i^{(T_i)}]$ where $E_i^{(j)} \in \mathbb{R}^d$ represents the embedding of the j^{th} token in the sequence, defined in equation (8). For downstream tasks, these embeddings are aggregated to produced a fixed-dimensional representation for mean pooling.

$$H_i = \frac{1}{T_i}\prod_{j=1}^{T_i} E_i^{(j)} \tag{8}$$

where $H_i \in \mathbb{R}^d$ is aggregated representation of the i^{th} review.

The resulting vector captures the overall semantic meaning of the sequence and is used as input for downstream tasks, such as classification or regression, as model working defined in Figure 4. (Kumari, 2023) This step ensures that the model produces a compact and meaningful representation of the entire input sequence.

Figure 4 Architecture of RoBERTa model (see online version for colours)



4.5 Dataset

The agriculture product review dataset includes 5K product reviews from users who assessed agricultural product quality on e-commerce sites. The dataset covers five product categories: Our data collection includes 1,000 product reviews from these categories, briefly defined in Table 2: fruits, vegetables, grains, dairy products, and spices. The reviews have sentiment scores assigned to them (positive or negative or neutral) and the reviewers detail their feedback on product quality elements including freshness, packaging, taste and purchase price value. The system includes price (USD) specifications and quality impact metrics to show product pricing and user ratings. This data is perfect for analysing customer emotions, assessing product quality, and learning what buyers want in agricultural online stores.

4.6 Performance evaluation measures

The performance of the proposed model is evaluated using standard classification metric. Let $Y = \{y_1, y_2, \ldots, y_N\}$ denotes the true labels and $\dot{Y} = \{\dot{y}_1, \dot{y}_2, \ldots, \dot{y}_N\}$ denote the predicted labels. Accuracy measures the proportion of correctly predicted instances out of the total instances, providing an overall assessment of model performance, computes as $Accuracy = \frac{1}{N}\sum_{t=1}^{N}(y_i = y_i')$. Precision quantifies the ratio of true positive predictions to all positive predictions, reflecting the model's ability to avoid false positives as showing $\frac{TP}{TP+FP}$. Recall calculates the ratio of true positive predictions to all actual positives, indicating the model's ability to identify all relevant instances, defined as $\frac{TP}{TP+FN}$. The F1-score is the harmonic means of precision and recall,

$2 \cdot \dfrac{Precision \cdot Recall}{Precision + Recall}$, balancing the two metrics to provide a single measure of model performance, especially useful in imbalanced datasets (Ishfaq et al., 2025). These metrics collectively evaluate the effectiveness of the model in classification tasks, based on dataset for each quality class $q_k \in \{q_1, q_2, \ldots, q_k\}$ and averaged to provide a comprehensive evaluation of the model's performance (Dong et al., 2024).

Table 2 Dataset labels description

Sr. no	*Product*	*Quality impact*	*Sentiment*
1	Fruits	Excellent taste	Positive
2	Grains	Nutrient rich	Negative
3	Vegetables	Okay packaging	Neutral
4	Daily products	High-quality packaging	-
5	Spices	Acceptable taste	-

5 Results and discussion

5.1 Descriptive analysis

These visualisations help to understand the distribution of various types of agricultural products and sentiment ratings for each of them. Figure 5 reveals the number of observations on different types of agricultural products including grains, vegetables, fruits, etc. It demonstrates an even distribution of the data collected and similarly low numbers for each type of product. This uniform distribution is crucial since it helps reduce bias on a specific product type in the subsequent analyses that utilise this dataset. Figure 6 zooms into the sentiment distribution for each of the product type dividing them into Positive, Neutral, and Negative rating. Regarding different products, it is found that the proportion of positive ratings is the highest, and then the neutral and the negative ones. This rise indicates that customers have a positive attitude towards most of these agricultural products. But there are differences in minor notes depending on the type of product. For example, Fruits and Spice seem to have moderately more neutral ratings compared to other products while negative ratings are always low. Such differences could be indicative of differences in overall customer experiences or perception of certain products. These visualisations highlight two important aspects of the dataset: the presence of all types of products and a positive sentiment trend in most cases. These results therefore offer a good foundation for carrying out subsequent analysis to find finer and more specific information regarding consumer preferences and product performance.

This kind of a visual representation is actually useful in understanding the linguistic characteristics of the user reviews pertaining to agricultural products. The word cloud in Figure 7 is the Word Cloud of User Reviews and includes some keywords, some of which appear bigger than others. As seen in the words on Figure 1, the most pinned terms include expectations, spices, dairy products, grains, purchased, fruits. These terms call attention to the fact that the users discuss the products with emphasis on several categories of products, more specifically, dairy products and grains. However, two noteworthy words are evident, one is 'expectations', which indicates that these customers have some expectations as to the quality of such products. This mean that often the description used are positive such as high quality, good and impressed which will make users give a positive sentiment regarding the products.

Figure 5 Distribution of product categories (see online version for colours)

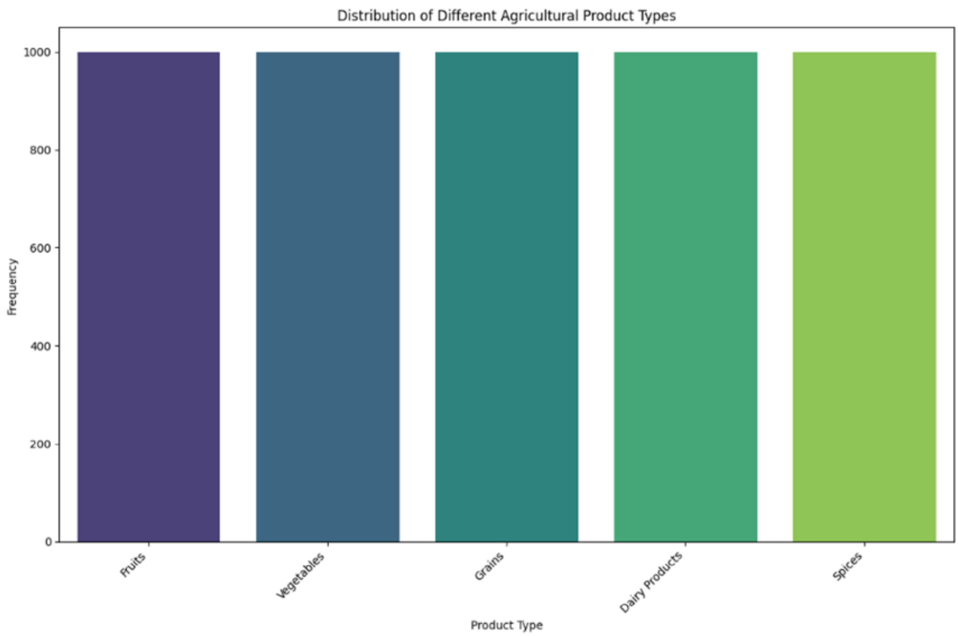


Figure 6 Product tags distribution by rating (see online version for colours)

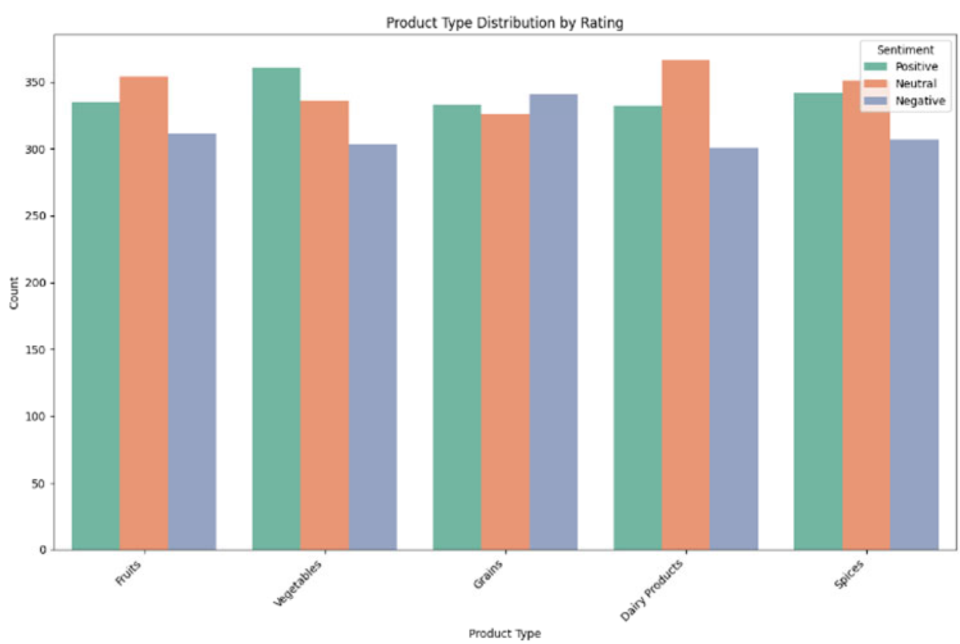


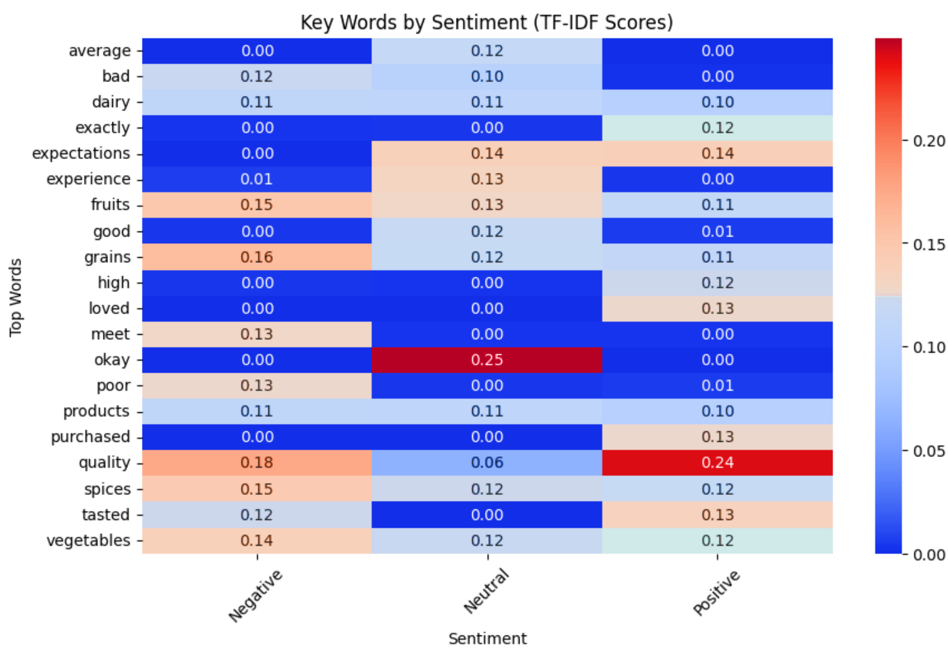
Figure 7 Wordcloud of users' review (see online version for colours)

In the three smaller word clouds for positive and negative polarity in Figure 8, the sentiment distribution between positive, negative, and neutral can be better observed. For the positive words: high quality, impressed, loved, exactly and so on, it means that many customers are satisfied with what they bought. Positive feedback more or less concerns factors such as the novelty of the products, their quality, and the clarity of product descriptions provided. On the other hand, the word cloud made from the neutral words concerns words like 'average quality', 'nothing special,' which point out that some users have neutral attitude and they did not note something positive or negative. The bad feeling in the negative word cloud contains words like; bad, poor quality, disappointing, and nothing extra-ordinary, in a way expressing the anger of customers who were let down by their experience. For instance, use of the words overpriced, received poor imply that the price offered was too high, quality of goods and the general condition of the received goods were not satisfactory. Collectively, these word clouds reveal that major products are linked to high word of mouth and general customer satisfaction, with a majority of the online opinions positive. However, it reveals negativity and neutrality meaning there are areas that can benefit from improvements in product quality, or management of customer expectations.

Figure 9 discussed above titled 'Keywords by sentiment (TF-IDF Scores)' is the matrix of sentiment and importance of key words. The intensity of colour in the matrix is pre-defined on the basis of the TF-IDF scores as shown in the next figure to highlight only the terms which have maximum importance in each sentimental class. Of all the positive sentiment words, there is no doubt that OK and ok have the highest TF-IDF score of 0.25 each and many App Store reviews do contain vernacular like the one below which portrays satisfaction or approval. At the same time, words such as poor and products stand out in the negative sentiment column as customers have used them when complaining. Other terms such as grains, fruts and dairy occur with moderate frequency across all groups of sentiment terms shown above, which indicate their frequent occurrence in the text, but which are not categorically associated with any particular sentiment.

Figure 8 Wordcloud of sentiment analysis based on reviews (see online version for colours)

Figure 9 Keywords by sentiments based on TF-IDF score (see online version for colours)

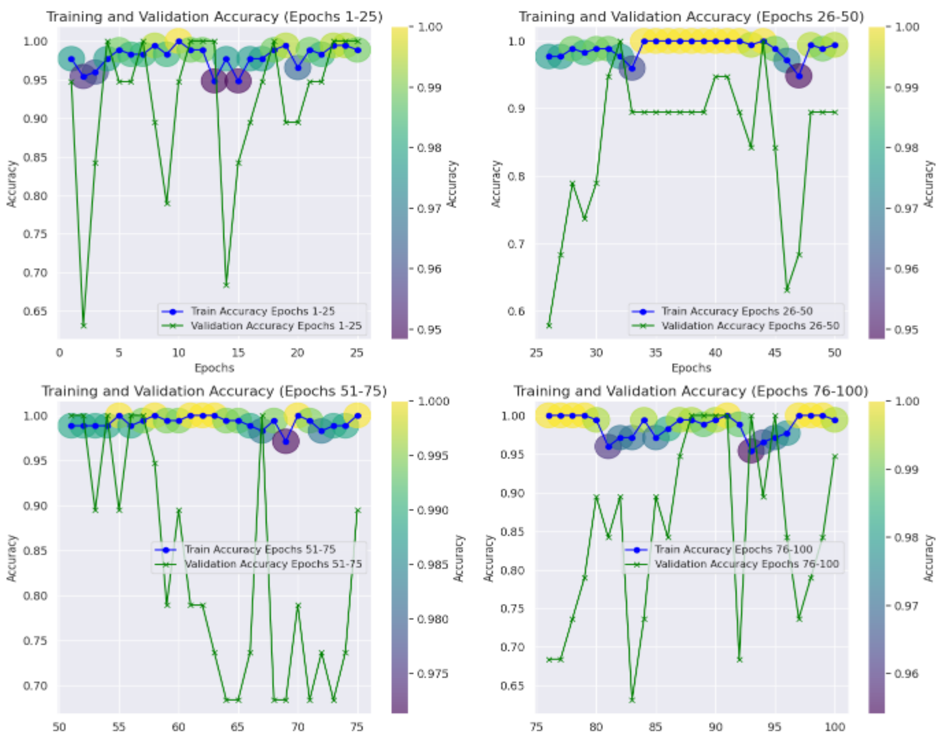


5.2 *Predictive analysis*

The fine-tuning of both BERT and RoBERTa models for the quality evaluation of agricultural products in e-commerce based on user reviews yielded very good results that give us a clear understanding of the capabilities of transformer- based models in sentiment analysis.

The BERT model test result shows that it was able to classify 96% of the reviews under different categories suggesting that the BERT model correctly marked 96% of the reviews. Accuracy has a problem in terms of lack of resolution as to how accurate the model performs relative to each type of error, especially when the specified dataset may have an uneven distribution of classes. However, the high level of accuracy indicates that the model meets the competence of the evaluation of positive, neutral, and negative sentiments in the user reviews and can be used as a suitable tool for quality assessment of products in the e-commerce environment. The precision of 89% means, out of all the times that BERT guessed a positive sentiment or quality review was correct 89% of the time. This precision value predicts true positive and but at the same time it shows one of the major problems inherent in this approach. The false positive rate regarding a neutral or negative account being identified as a positive one is 10%, as training and validation accuracy analysis is shown in Figure 10. In the scenario of executing e-commerce product review, a false positive could in some way have negative effects; for case, rating a low-quality product as a high-quality product may lead to deceiving customers. However, with 94% of recall, BERT achieves what it aims, pointing out 94% of true positives, while labelling only 6% of positive tone as negatives. This is important since it minimises cases in which negative feedback is missing, while positive feedback may persuade customer to purchase.

Figure 10 Accuracy analysis of BERT model over epochs (see online version for colours)



Moreover, RoBERTa achieves the highest accuracy of 99% to classify users' review analysis based on agriculture product. Also, the value of the 96% recall RoBERTa obtained implies that it can capture nearly all true positive instances and suffers from a 4% mismatch rate. This indicates that RoBERTa retains an advantage over BERT in its capability of positive sentiment classification and may attribute this sensitivity to improvement in the model's acculturation into client reviews. Table 3 highlights the applied hyperparameters for setting experimental setup of RoBERTa Model. In practical terms, this ensures that RoBERTa does not miss most positive product assessments in customer feedback, which also means high-grade products will not be permanently shelved due to unpleasant review of users. In more detail, Table 4 analyses the results by each of the metrics and considers their potential for evaluating overall product quality. RoBERTa scores 98% on the F1 score, which is even more optimal in balanced precision and recall, which makes the model more robust and accurate when reviewing products and performing evaluations, as accuracy analysis over epochs shown in Figure 11.

Table 3 Applied hyperparameters of proposed model

Hyperparameter	*Description*	*Value*
Model	Pre-trained language model used for fine-tuning	RoBERTa-large
Learning rate	Step size for gradient descent	2e-5
Batch sequence	Number of samples processed before model update	16 (train), 32 (eval)
Max sequence length	Maximum number of tokens per input sequence	128 or 512
Epochs	Number of complete passes through the dataset	100
Optimiser	Optimisation algorithm	AdamW
Warmup steps	Number of steps for learning rate warmup	500
Weight decay	Regularisation technique to prevent overfitting	0.01
Evaluation strategy	Determines when to evaluate during training	Steps or epoch
Gradient accumulation	Accumulate gradients over several batches before updating weights	2
Dropout rate	Fraction of neurons randomly set to zero during training	0.1
Seed	Random seed for reproducibility	42
Loss function	Loss function used for classification tasks	CrossEntropyLoss
Scheduler	Learning rate adjustment scheduler	Linear Scheduler

Such results indicate an improvement of not only RoBERTa but also BERT in terms of F1 score and it also proves its dominance to be more stable across categories of products and reviews. This model-oriented approach is ideal in e-commerce because it tackles the challenges of dealing with large datasets with diverse sentiments, while providing the best assessments in terms of precision and recall.

Both BERT and RoBERTa are highly successful in predicting product quality from customers' reviews, as shown in Figure 12. In any e-commerce transaction, the ability to understand user reviews and predict quality is vital for many reasons including determining what products a consumer buys, what items are recommended to the customer, and what quality parameters should be set for the products.

In this context, precision is of tremendous importance since businesses need to focus on the right items to sell and recommend to consumers. With RoBERTa precision rated at 92% and BERT at 89%, the chances of poor-quality products being recommended are extremely low. This is extremely crucial because with these methods in place, consumers will only get recommendations for products they know they can trust. Lowering too much precision would result in a high false positive rate which would add a stigma to the platform and result in terrible customer service encounters. Higher precision further helps businesses in cutting out these risk avenues.

Figure 11 Accuracy Analysis of RoBERTa model over epochs (see online version for colours)

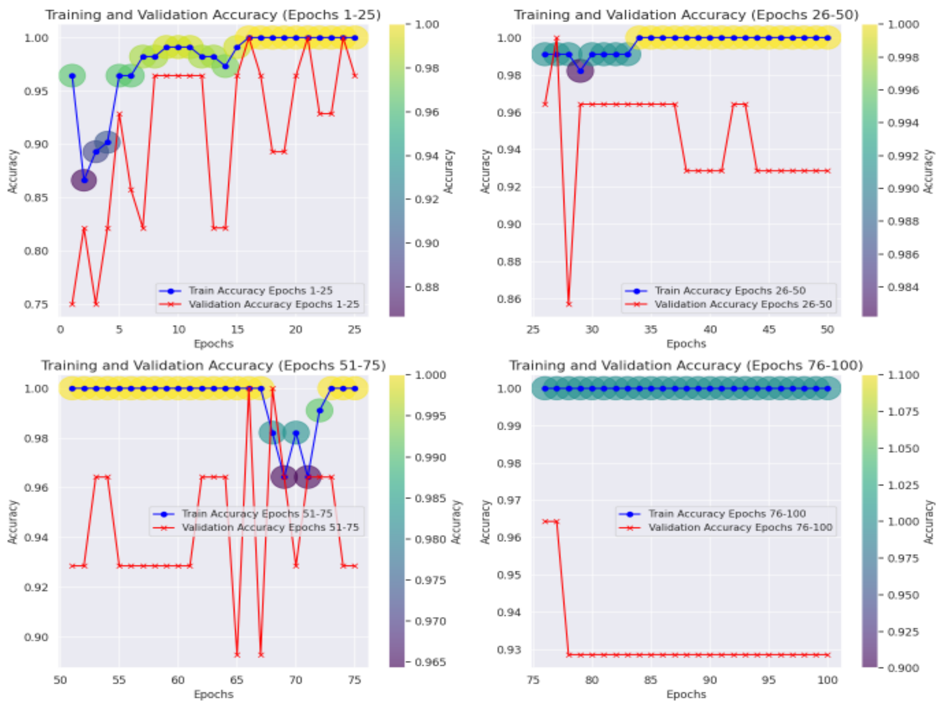


Figure 12 Results analysis of applied model (see online version for colours)

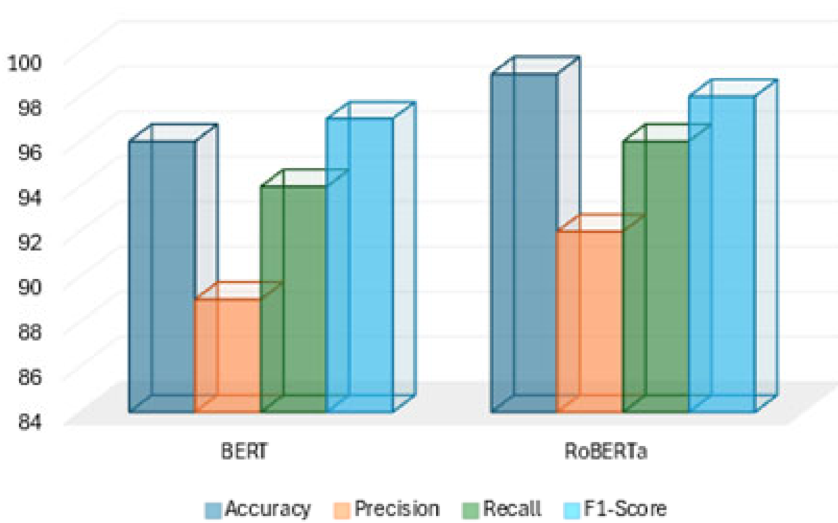


Table 4 Results with applied models

	Accuracy	*Precision*	*Recall*	*F1-Score*
BERT	96	89	94	97
RoBERTa	99	92	96	98

5.3 *Implications of model performance*

The trust consumers have for an e-commerce platform has everything to do with how the platform utilises user reviews and product ratings. Both BERT and RoBERTa have exceptional prediction accuracy, however, one of them has proved to be more effective in areas where recommendations need to be extremely accurate: the latter. With its source material, RoBERTa is reliable for cutting-edge recommendations that adopt AI systems, following in Algorithm 1.

Users can easily be assisted in picking the most authentic Agri-products because they have already been scrutinised and picked up by real users. Besides, the ability of both models to extrapolate to various types of product categories and reviews means that there is potential for cross marketing within and outside the agriculture industry. The use of these transformer models enables e commerce platforms to design better quality assessment systems, eliminating overreliance of manual systems and increasing productivity. In addition, these models will go a long way in enhancing the consumer's experience because a long-term relationship will be built based on trust in the high quality of products recommended.

Algorithm: Proposed Model RoBERTa

Input: Pre-trained RoBERTa-large model, input sequence, labels

Output: Fine-tuned RoBERTa-Large model

For each x in Pre-trained RoBERTa model:

$RoBERTa\_Large[x] \leftarrow Loadd(PreTrainedModel)$

End for

For each x Input_Sequence:

$Tokenised[x] \leftarrow Tokenised(Input_{Sequence[x]})$

$Encoded[x] \leftarrow RoBERTa_{Large[x]}, encode(Tokenised[x])$

End for

For each x, y in Encoded:

$Q[x] \leftarrow Encoded[x] \,.\, W_q$

$K[y] \leftarrow Encoded[y] \,.\, W_k$

$V[x] \leftarrow Encoded[x] \,.\, W_v$

$Attention\_weights[x, y] \leftarrow softmax((Q[x]\ K[y]^T)/\ \sqrt{d_k}\,)$

$Output_{attention[x, y]} \leftarrow Attention\_weights[x, y] \,.\, V[x]$

$FF_{output[x, y]} \leftarrow ReLU(Output_{attention[x, y]} \,.\, W_L + b_L$

$Final_i Output_{[x, y]} \leftarrow FF_{output[x, y]} \,.\, W_L + b_L$

End for

For each x in Final_output:

$H[x] \leftarrow Final_x Output[x] + Encoded[x]$

$H_{normalised}[x] \leftarrow LayeNorm(H[x])$

End for
For each x in Task_Labels:
$Task_{output}[x] \leftarrow H_{normalized}[x]W_i^T task + b\_task$
End for
For each x in Task_output:
$Loss[x] \leftarrow CrossEntropyLoss(Task_{output[x]}, GroundTruth[x])$
End for
For each x in Loss:
$Gradient[x] \leftarrow \nabla Loss[x]$
$RoBERTa_{Large[x]} \leftarrow RoBERTa_{Large[x]} - Learning_{rate} \,.\, Gradients[x]$
End for
For each x in RoBERTa_Large:
$FineTuned_{Mode[x]} \leftarrow RoBERTa\_Large[x]$
End for

Table 5 Comparison of baseline (BERT) and proposed model (RoBERTa)

Feature	*BERT (baseline model)*	*RoBERTa (proposed model)*
Architecture	Transformer-based, bidirectional	Transformer-based, improved bidirectional
Pretraining data	16GB of text (BooksCorpus + Wikipedia)	160GB of text (CommonCrawl + BooksCorpus + OpenWebText + Wikipedia)
Training strategy	Next Sentence Prediction (NSP) + Masked Language Modeling (MLM)	Removed NSP, improved dynamic MLM
Training time	Shorter training time	Longer training with more data
Tokenisation	WordPiece Tokenisation	Byte-Pair Encoding (BPE)
Performance	Strong baseline for NLP tasks	Outperforms BERT in text classification and sentiment analysis
Fine-tuning efficiency	Requires more hyperparameter tuning	More robust with less fine-tuning

In conclusion, while BERT offers solid performance for quality assessment, RoBERTa is the more robust and reliable model, providing superior accuracy and a better balance between precision and recall, as display in Table 5. This makes RoBERTa the ideal model for e-commerce platforms focused on delivering high-quality product recommendations based on user feedback.

Table 5 provides a structured comparison to highlight why RoBERTa was chosen over BERT for the proposed model.

5.4 Comparison with existing studies

Table 5 evaluates different approaches to measure agricultural product quality through analysis of user input. The default parameter setup of RoBERTa produces the best results at 99% by effectively picking up detailed information from user reviews. The method outperforms alternative models by applying advanced text review-native NLP methods.

Studies of Chinese agricultural reviews using BiLSTM and CNN models with FastText and Word2Vec achieved 89% accuracy in quality assessment. The chosen methods demonstrated potential, but their standard embedding techniques prevented them from fully understanding advanced text relationships. The performance of SVM and NB models analysing crop and soil records ranges from 78% to 90%, reflecting their benefits and challenges in processing user-created agricultural documentation. BERT reached 91% accuracy by using FastText and Word2Vec embedding technology. These models adopt traditional approaches to modern transformer systems for better context analysis. The RoBERTa framework achieves improved results by using a purpose-built fine-tuned system that handles the unique linguistic features present in agricultural product reviews. The comparison shows that transformer-based models like RoBERTa provide essential benefits for quality evaluation work. Strong results validate the benefits of domain-specific tuning for pre-trained models that yield more precise feedback about users and their products.

Table 6 Comparison with existing studies

Ref	*Year*	*Model*	*Dataset*	*Features*	*Results (%)*
He et al.	2020	Bi-LSTM, CNN	Chinese products reviews	FastText, and Word2Vec	89
Benos et al.	2021	ANN, SVM	Crop dataset	Textual features	78
Cravero et al.	2022	SVM, DT, NB	Historical records of soil	Textual features	80
Alghazzawi et al.	2023	ERF-XGB	IMDB	Harris hawk optimisation	92
Bellar et al.	2024	BERT	Reviews of different products	FastText, and Word2Vec	91
Yuan et al.	2025	BERT	Chinese agriculture product user's review	Default	88
Proposed	2025	RoBERTa-Large	Agriculture product user's review	Fine-Tuned-Default Features	99

6 Conclusions and future work

In today's era of e-commerce, analysing user-generated reviews has become critical for assessing product quality and ensuring customer satisfaction. Leveraging advanced AI models, particularly fine-tuned masked language models such as BERT and RoBERTa substantially improve efficiency for the textual objective when handling large volume of textual data. The significance of this study lies in its application of state-of-the-art transformer-based DL models to the domain of agricultural product quality assessment, addressing the challenges of extracting insights from subjective customer reviews. It not only enhances the quality of a customer's decision making but also enhances the provision of high-quality goods by businessmen. Our findings demonstrate, transformer-based models outperform other models; among them, RoBERTa yielded the highest results according to all the estimated criteria – up to 99% of accuracy, Unlike BERT, for which similar performance is found, RoBERTa fine-tuning gives a deeper analysis of the customer's feeling and gives fewer errors in the prediction. These findings thus demonstrate the effectiveness of better language models in producing reasonable quality evaluations allowing for the problems of high variability and subjectivity of the feedback in certain contexts, such as e-commerce, to be solved successfully. A key limitation of this study is its reliance solely on textual data from customer reviews, which may overlook valuable insights from other modalities such as product images, videos, or structured metadata. Hence, future work can advance this research by employing other features along with text data including the product images or video reviews to achieve even higher accuracy of prediction. Furthermore, proposed domain-specific transformer models towards agricultural products Reviews could enhance the fine-grained sentiment analysis and share resources for relevant fields.

Funding

Discipline Co construction Project of Philosophy and Social Sciences in Guangdong Province's 13th Five Year Plan for 2020 (Project Code:GD20XJY30).

Youth Innovation Talent Project of Guangdong Provincial Department of Education (Project Code: 2022WQNCX144).

Declarations

All authors declare that they have no conflicts of interest.

References

Ahmad, W., Khan, H.U., Iqbal, T. and Iqbal, S. (2023) 'Attention-based multi-channel gated recurrent neural networks: a novel feature-centric approach for aspect-based sentiment classification', *IEEE Access*, January, Vol. 11, No. 6, pp.54408–54427.

Ahmed, M., Khan, H.U. and Munir, E.U. (2023) 'Conversational AI: an explication of few-shot learning problem in transformers-based chatbot systems', *IEEE Transactions on Computational Social Systems*, Vol. 11, No. 2, pp.1–19.

Alghazzawi, D., Alquraishee, A., Badri, S. and Hasan, S.S. (2023) 'ERF-XGB: ensemble random forest-based xg boost for accurate prediction and classification of e-commerce product review', *Sustainability*, April, Vol. 15, No. 9, pp.7076–7076, DOI: https://doi.org/10.3390/su15097076.

Ali, H., Hashmi, E., Yildirim, S.Y. and Shaikh, S. (2024) 'Analyzing Amazon products sentiment: a comparative study of machine and deep learning, and transformer-based techniques', *Electronics*, January, Vol. 13, No. 7, p.1305, DOI: https://doi.org/10.3390/electronics 13071305.

Alsaedi, T., Rana, M.R.R., Nawaz, A. Raza, A. and Alahmadi, A. (2024) 'Sentiment mining in e-commerce', *International Journal of Electrical and Computer Engineering Systems*, September, Vol. 15, No. 8, pp.641–650, DOI: https://doi.org/10.32985/ijeces.15.8.2.

Alzahrani, M.E., Aldhyani, T.H.H., Alsubari, S.N., Althobaiti, M.M. and Fahad, A. (2022) 'Developing an intelligent system with deep learning algorithms for sentiment analysis of e-commerce product reviews', *Computational Intelligence and Neuroscience*, May, Vol. 2022, No. 1, pp.1–10.

Arobi, R., Rijon, Hassan, R., Haque, N., Safkat, Sakir, T.I.M. and Safoan, S. (2022) 'Sentiment analysis on e-commerce based product reviews using machine learning algorithms', *Bracu.ac.bd*, DOI: https://doi.org/ID%2017101330.

Bellar, O., Baina, A. and Ballafkih, M. (2024) 'Sentiment analysis: predicting product reviews for e-commerce recommendations using deep learning and transformers', *Mathematics*, August, Vol. 12, No. 15, pp.2403–2403.

Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D. and Bochtis, D. (2021) 'Machine learning in agriculture: a comprehensive updated review', *Sensors*, January, Vol. 21, No. 11, p.3758, DOI: https://doi.org/10.3390/s21113758.

Cao, Y., Sun, Z., Li, L. and Mo, W. (2022) 'A study of sentiment analysis algorithms for agricultural product reviews based on improved BERT model', *Symmetry*, Vol. 14, No. 8, p.1604.

Cravero, A., Pardo, S., Sepúlveda, S. and Muñoz, L. (2022) 'Challenges to use machine learning in agricultural big data: a systematic literature review', *Agronomy*, March, Vol. 12, No. 3, p.748, DOI: https://doi.org/10.3390/agronomy12030748.

Daza, A., González, D., Aguilar, S., Filomeno, W. and Chauca, E. (2024) 'Sentiment analysis on e-commerce product reviews using machine learning and deep learning algorithms: a bibliometric analysis and systematic literature review, challenges and future works', *International Journal of Information Management Data Insights*, November, Vol. 4, No. 2, pp.100267–100267.

Dong, Y., Zhang, Y., Wu, G., Chen, X., Tu, L. and Fan, G. (2024) 'Sentiment classification method of fresh agricultural product reviews based on semantic and emotional optimisation', *International Journal of Intelligent Information and Database Systems*, December, Vol. 17, No. 1, pp.32–56, DOI: https://doi.org/10.1504/ijiids.2025.143485.

Fidyanti, S., Hadju, N. and Jayadi, R. (2021) 'Sentiment analysis of indonesian e-commerce product reviews using support vector machine based term frequency inverse document frequency', Vol. 99, No. 17 [online] https://www.jatit.org/volumes/Vol99No17/9Vol99No17.pdf.

He, Z., Yong, Y., Guo, F. and Zhang, X. (2020) 'Sentiment analysis of agricultural product ecommerce review data based on deep learning', *2020 International Conference on Internet of Things and Intelligent Applications (ITIA)*, Zhenjiang, China, Vol. 2020, No. 9, pp.1–7.

Huang, H., Zavareh, A.A. and Mustafa, M.B. (2023) 'Sentiment analysis in e-commerce platforms: a review of current techniques and future directions', *IEEE Access*, Vol. 11, No. 8, pp.90367–90382.

Ishfaq, U., Khan, H.U. and Shabbir, D. (2025) 'Exploring the role of sentiment analysis with network and temporal features for finding influential users in social media platforms', *Social Network Analysis and Mining*, January, Vol. 14, No. 1, DOI: https://doi.org/10.1007/s13278-024-01396-6.

Jahan, H. et al. (2022) 'Sentiment analysis of e-commerce consumer based on product delivery time using machine learning', *Lecture Notes on Data Engineering and Communications Technologies*, January, pp.649–661, DOI: https://doi.org/10.1007/978-981-19-2347-0_51.

Khrais, L.T. (2020) 'Role of artificial intelligence in shaping consumer demand in e-commerce', *Future Internet*, December, Vol. 12, No. 12, p.226.

Kumari, K. (2023) 'RoBERTa: a modified BERT model for NLP', *Comet*, 22 September [online] https://www.comet.com/site/blog/roberta-a-modified-bert-model-for-nlp/ (accessed 9 January 2025).

Lin, X. (2020) 'Sentiment analysis of e-commerce customer reviews based on natural language processing', *Proceedings of the 2020 2nd International Conference on Big Data and Artificial Intelligence*, April.

Liu, Y. (2024) 'Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax', *Journal of Corrosion and Materials*, December, Vol. 48, No. 1, pp.42–57.

Naqvi, S.M.M.R., Batool, S., Ahmed, M., Khan, H.U. and Shahid, M.A. (2023) 'A novel approach for building domain-specific chatbots by exploring sentence transformers-based encoding', *IEEE Xplore*, 1 October [online] https://ieeexplore.ieee.org/document/10335884 (accessed 15 March 2024).

Nawaz, Z., Zhao, C., Nawaz, F., Safeer, A.A. and Irshad, W. (2021) 'Role of artificial neural networks techniques in development of market intelligence: a study of sentiment analysis of eWOM of a women's clothing company', *Journal of Theoretical and Applied Electronic Commerce Research*, August, Vol. 16, No. 5, pp.1862–1876.

Naz, A., Khan, H.U., Alesawi, S., Abouola, O.I., Daud, A. and Ramzan, M. (2024) 'AI knows you: deep learning model for prediction of extroversion personality trait', *IEEE Access*, Vol. 12, pp.159152–159175, DOI: 10.1109/ACCESS.2024.3486578.

Noor, A. and Islam, M. (2019) 'Sentiment analysis for women's e-commerce reviews using machine learning algorithms', *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, July.

Rong, L., Weibai, Z. and Debo, H. (2021) 'Sentiment analysis of ecommerce product review data based on deep learning', *IEEE Xplore*, 1 June, Vol. 4, No. 7.

Shang, F., Shi, J., Shi, Y. and Zhou, S. (2014) 'Enhancing e-commerce recommendation systems with deep learning-based sentiment analysis of user reviews', *International Journal of Engineering and Management Research*, Vol. 14, No. 4, pp.19–34.

Singh, S.N. and Sarraf, T. (2020) 'Sentiment analysis of a product based on user reviews using random forests algorithm', *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, January, DOI: https://doi.org/10.1109/confluence47617.2020.9058128.

Talha, M.M., Khan, H.U., Iqbal, S., Alghobiri, M., Iqbal, T. and Fayyaz, M. (2023) 'Deep learning in news recommender systems: a comprehensive survey, challenges and future trends', *Neurocomputing*, December, Vol. 562, pp.126881–126881.

Taneja, K., Vashishtha, J. and Ratnoo, S. (2024) 'Transformer based unsupervised learning approach for imbalanced text sentiment analysis of e-commerce reviews', *Procedia Computer Science*, January, Vol. 235, pp.2318–2331, DOI: https://doi.org/10.1016/j.procs.2024.04.220.

Vanneschi, L., Horn, D.M., Castelli, M. and Popovič, A. (2018) 'An artificial intelligence system for predicting customer default in e-commerce', *Expert Systems with Applications*, August, Vol. 104, No. 8, pp.1–21.

Xiang, Y., Yu, H., Gong, Y., Huo, S. and Zhu, M. (2024) *Text Understanding and Generation using Transformer Models for Intelligent E-Commerce Recommendations*, ArXiv (Cornell University), Vol. 4, No. 3, pp.170–179, https://doi.org/10.1117/12.3034062.

Xu, F., Pan, Z. and Xia, R. (2020) 'E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework', *Information Processing & Management*, February, Vol. 57, No. 5, p.102221.

Yang, L., Li, Y., Wang, J. and Sherratt, R.S. (2020) 'Sentiment analysis for e-commerce product reviews in chinese based on sentiment lexicon and deep learning', *IEEE Access*, Vol. 8, No. 1, pp.23522–23530.

Yuan, L., Liu, H., Fu, F., Liu, Y., Zuo, X. and Li, L. (2025) 'Study of Zhejiang tangerine e-commerce reviews based on natural language processing', *Horticulturae*, February, Vol. 11, No. 2, p.151, DOI: https://doi.org/10.3390/horticulturae11020151.