

Sentiment Analysis on Malaysian Company Reviews from Glassdoor

Nur Azmina Mohamad Zamani^{1*}, Ahmad Muhyiddin Yusof², Nur Fatihah
Syafika Ahmad Firdauz³

¹*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*

²*Academy of Language Studies, Pusat Asasi, Universiti Teknologi MARA Selangor Branch, Dengkil Campus, 43800 Dengkil, Selangor, Malaysia*

³*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Perak Branch, Tapah Campus, 35400 Tapah Road, Perak, Malaysia*

ARTICLE INFO

Article history:

Received 16 April 2026

Revised 28 May 2026

Accepted 14 June 2026

Online first

Published 5 October 2026

Keywords:

Aspect-based Sentiment Analysis

Natural Language Processing

Malaysian Company Reviews

Glassdoor

GPT-3.5-turbo

DOI:

10.24191/mij.v7i2.10899

ABSTRACT

The growing number of job seekers has increased reliance on company reviews for workplace insights. However, the unstructured nature of these reviews makes systematic interpretation challenging. This study performs Aspect-Based Sentiment Analysis (ABSA) on 8,412 Glassdoor reviews (2012–2024) across 10 major Malaysian companies using GPT-3.5-turbo. The system classifies sentiment (positive, negative, neutral) across five aspects: Work-Life Balance, Culture & Values, Benefits & Compensation, Career Opportunities, and Senior Management. Using three core ABSA tasks—Aspect Term Extraction (ATE), Aspect Detection and Classification (ADC), and Aspect Sentiment Classification (ASC)—a prompt-engineered few-shot learning approach was developed with 15 contextual JSON examples and a zero-temperature setting. Insights are visualized via an interactive web-based dashboard. Evaluation revealed average Aspect Accuracy of 84.14%, Sentiment Accuracy of 94.19%, Component Accuracy of 66.73%, and Exact Match Accuracy of 64.62%. Comparative benchmarking against VADER, ATAE-LSTM, and fine-tuned BERT models demonstrated superior sentiment classification accuracy and exact match performance without model retraining. System Usability Scale (SUS) testing with 30 users yielded an "Excellent" rating of 81.25. Overall, the system transforms unstructured reviews into actionable insights to help job seekers identify aligned employment opportunities.

^{1*} Corresponding author. *E-mail address:* azmina@uitm.edu.my

1. INTRODUCTION

Glassdoor is a site for job seekers and recruiters worldwide. There are millions of job listings with reviews and ratings available on the website (Galiano-Coronil et al., 2024). The company reviews can be an influential component in determining job seekers' decisions in selecting a job suited to their needs. Employee-generated information offers detailed insights into workplace realities, frequently exceeding official company narratives in perceived authenticity (Kaur & Dubey, 2014).

However, the vast amount of unstructured textual reviews makes analysis difficult to perform. Traditional sentiment analysis usually produces a single polarity label for an entire document or piece of content. This technique tends to overlook the multidimensional nature of workplace feedback since a single review can simultaneously praise compensation and criticise management practices at the same time. Due to the limitation of a single-labelled polarity on a document, Aspect-Based Sentiment Analysis (ABSA) is implemented to identify the specific aspects contained in the texts while determining the sentiment of the particular content (Ahmed et al., 2023).

Currently, there is still limited research on company review analytics in Malaysia, regardless of the growing digital transparency in recruitment ecosystems. Therefore, this study proposes a GPT-driven ABSA system adapted to Malaysian company reviews. The system focuses on five core workplace aspects that often involve employee satisfaction (positive, negative, and neutral). The five aspects are Work-Life Balance, Culture & Values, Benefits & Compensation, Career Opportunities, and Senior Management.

The main contributions of this work are (1) establishing a structured dataset of 8,412 employee reviews across 10 top Malaysian public and private enterprises; (2) implementing a prompt-engineered few-shot GPT-3.5 Turbo ABSA pipeline capable of simultaneous ATE, ACD, and ASC; (3) benchmarking GPT performance against conventional lexicon, LSTM, and fine-tuned BERT models; and (4) developing and empirically evaluating an interactive web application for visual organizational analytics and usability. This paper is organised as follows: Section 2 will present the literature review, Section 3 explains the methodology, the results and discussion are described in Section 4, and the conclusion in Section 5.

2. LITERATURE REVIEW

The emergence of online company review platforms such as Glassdoor has transformed employee feedback into an open-access information source, exposing organization work environment and management to all users. The platform contains experiential insights that can give job seekers some ideas about the nature of the workplace and allows the organisation to enhance their management. As a result, company reviews function as personal testimonials and a data source that reveals systematic patterns in organisational strengths and weaknesses. Thus, the application of computational techniques such as Aspect-based Sentiment Analysis is useful to understand the structure of the reviews.

2.1 Company Review and Workplace Aspects

Company reviews provide valuable information on employee satisfaction and workplace dimensions. In the aspect of work-life balance, an optimal balance is achieved when the employees can function in both professional and personal roles effectively by minimising conflicts (Tijani et al., 2022). On the other hand, inflexible schedules and unavoidable overtime can cause burnout, lower productivity, and stronger intentions to leave the organisation among employees (Tijani et al., 2022).

Another important aspect in the workplace is the career advancement opportunities. This aspect is important to shape employee motivation and long-term retention. Professional growth improves work

engagement and strengthens affective commitment (Zhu & Song, 2022). Hence, initiatives such as mentorship programs, workplace training, and leadership development pathways enable employees to seek progression. In contrast, limited prospects for career development can lead to an increase in turnover intentions and reduced engagement. By prioritising continuous learning and structured career pathways, organisations can improve employee satisfaction while cultivating a competent workforce that supports strategic objectives.

Compensation is a fundamental determinant of employee motivation and work performance, as it directly influences individuals' quality of life, social status, and overall job satisfaction. According to Ullah et al. (2024), employees who receive competitive remuneration tend to demonstrate higher levels of commitment and productivity. A strategically designed compensation and benefits framework that reflects market standards, performance evaluations, and organisational objectives can improve satisfaction while simultaneously attracting and retaining talented employees (Adesina & Egbuta, 2025). This can positively reflect organisational values. Strong leadership is a key factor in organisational culture and employee morale. Practices such as consistent feedback, transparent communication, and clear acknowledgement of employee achievements contribute to greater trust in management (Raza et al., 2023). Conversely, ineffective leadership behaviours, including excessive control and insufficient accountability, often lead to unfavourable perceptions and reduced job satisfaction (Modise, 2023). Organisations led by capable teams that encourage teamwork and innovation are therefore more likely to achieve sustained success and maintain high levels of employee engagement.

A well-established organisational culture that aligns with employees' personal values contributes significantly to higher job satisfaction and improved performance. Vijayakumar and Rajagopal (2023) emphasise that having a constructive work environment can lead to minimal aggression and harassment by ensuring safety, fairness, regulatory compliance, and the protection of employee rights. Moreover, diversity, equity, and inclusion are important cultures that should be practised in organisations. This is because a toxic environment can result in disengagement and turnover.

2.2 Aspect-Based Sentiment Analysis (ABSA)

Aspect-Based Sentiment Analysis (ABSA) is an advanced subfield of sentiment analysis that goes beyond identifying the overall polarity of a text by determining sentiment toward specific attributes or components mentioned within it. Unlike traditional document-level approaches that assign a single positive or negative label to an entire review, ABSA decomposes text into finer-grained elements known as aspects.

The ABSA process typically involves two main tasks: aspect extraction and sentiment classification. Aspect extraction identifies relevant features or topics discussed in the text, either through predefined categories or automated techniques such as keyword extraction, topic modelling, or sequence labelling methods. Once aspects are detected, sentiment polarity is assigned to each identified aspect using lexicon-based scoring, supervised machine learning algorithms, or deep learning architectures. This two-stage process ensures that sentiment is accurately linked to the appropriate feature, reducing ambiguity in interpretation. Fig. 1 shows the conceptual framework of ABSA on employee review text.

Methodologically, ABSA can be implemented using rule-based systems, classical machine learning techniques, or neural network models. Rule-based and lexicon-driven approaches rely on sentiment dictionaries and syntactic patterns to associate opinion words with specific aspects. While these methods offer interpretability and lower computational requirements, it still struggles with contextual subtleties such as sarcasm or implicit sentiment. Machine learning-based ABSA models, including Support Vector Machines (SVM), Conditional Random Fields (CRF), and Naïve Bayes classifiers, learn aspect-sentiment relationships from labelled datasets. More recently, deep learning models such as Long Short-Term

Memory (LSTM) networks and transformer-based architectures have demonstrated improved capability in capturing contextual dependencies and semantic relationships within text.

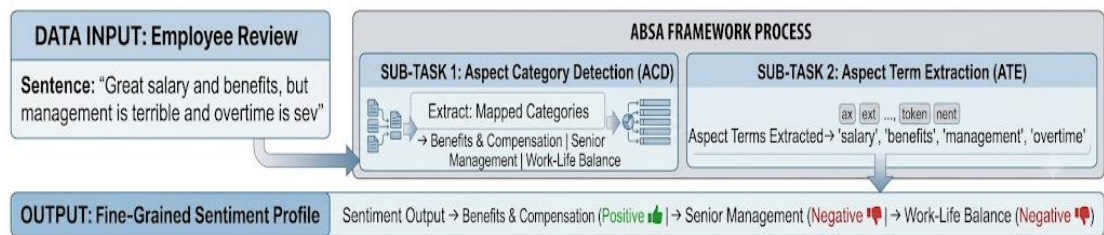


Fig. 1. Conceptual Framework of Aspect-Based Sentiment Analysis (ABSA) on Employee Review Text

Source: Authors' own illustration

In organisational analytics, ABSA provides strategic value by enabling decisionmakers to identify precisely which workplace factors drive satisfaction or dissatisfaction. Rather than relying solely on overall sentiment scores, managers can evaluate patterns across distinct organisational dimensions and prioritize targeted interventions. For instance, consistently negative sentiment associated with career advancement may signal the need for structured development programs, while positive sentiment regarding culture may reinforce current practices. Therefore, ABSA serves as a significant analytical framework for transforming unstructured employee feedback into actionable organisational intelligence, supporting evidence-based improvements and sustainable workforce management. Fig. 1 shows an example of aspect-sentiment classification.

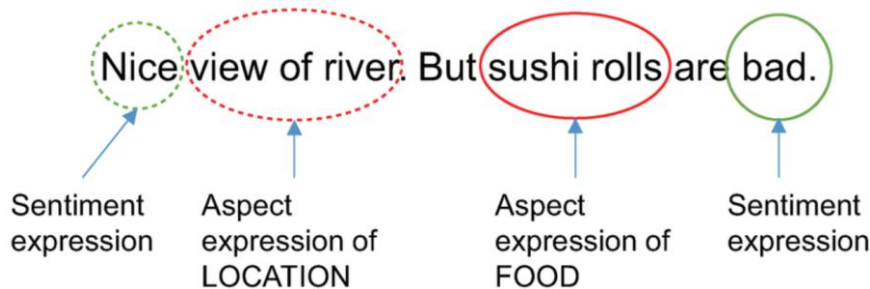


Fig. 2. Example of Aspect-based Sentiment Classification (Noh et al., 2019)

Source: Authors' own illustration

Based on Fig. 2, the study by Noh et al. (2019) applied Aspect-Based Sentiment Analysis (ABSA) and highlighted that sentiment expressions usually occur near their corresponding aspect expressions. For example, a restaurant review may describe the location as "view of river" with a positive word like "nice" nearby, and the food as "sushi rolls" with a negative term such as "bad" in close proximity. Although leveraging such positional information can enhance ABSA methods, corpora with word-level annotations for aspects and sentiments are scarce. Most available datasets provide only text-level sentiment annotations, and constructing word-level annotated corpora is a labour-intensive and time-consuming task. There are three main tasks involved in performing ABSA, which are Aspect Term Extraction (ATE), Aspect Detection and Classification (ADC), and Aspect Sentiment Classification (ASC).

Aspect Term Extraction (ATE) involves identifying explicit aspect expressions present in a text. It is commonly considered as a Named Entity Recognition (NER) task, where aspects are detected through sequence labelling (Zhou et al., 2023). Typically, these are nouns or noun phrases that indicate the specific features or topics the writer is addressing. For instance, in the sentence “The user interface is outdated and slow,” both “user interface” and possibly “speed” can be recognised as aspect terms. At this stage, ATE focuses only on determining what is being discussed, without evaluating the writer’s sentiment toward it. Fig. 3 shows an example of ATE.

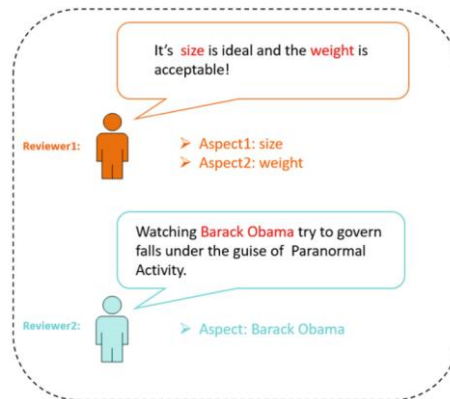


Fig. 3. Example of Aspect Term Extraction (ATE) (Zhang & Yang, 2023)

Source: Authors' own illustration

Fig. 3 shows Aspect Term Extraction (ATE) examples from two reviewers. Reviewer 1 comments on a product, highlighting “size” and “weight” as separate aspect terms. Reviewer 2 reviews a movie, identifying “Barack Obama” as the aspect term, showing that aspects can also include proper nouns. The image demonstrates how ATE extracts specific targets from text, regardless of whether they are product features, characteristics, or entities.

While Aspect Term Extraction (ATE) identifies explicit aspect expressions in text, Aspect Category Detection (ACD) extends this by mapping sentences to a predefined set of broader and more abstract categories, even if these categories are not explicitly mentioned. For instance, the sentence “I’m always expected to work overtime” does not directly state “Work-Life Balance,” yet it implicitly relates to that concept. ACD is particularly valuable in domains where a fixed set of categories is established, such as common topics in employee or product reviews. This task depends on interpreting the contextual meaning of text and accurately linking it to the relevant category, requiring both linguistic understanding and domain knowledge, since the connection between expressions and categories is often indirect (Chebolu et al., 2022). In structured review systems or surveys, ACD enables analysts to quantify feedback consistently across categories, facilitating comparison and aggregation of sentiments.

Aspect Sentiment Classification (ASC) involves identifying the sentiment (i.e., positive, negative, or neutral) that associated with a certain aspect or category within a text. Once an aspect has been recognised or categorised, ASC assesses the writer’s opinion regarding the specific aspect. For example, in the sentence “The promotion process is unfair,” the sentiment toward the aspect “promotion process” is negative, whereas in “The training sessions were incredibly helpful,” the sentiment associated with “training sessions” is positive. Accurate ASC requires examining both the aspect and its surrounding context to interpret sentiment correctly (Brauwiers & Frasincar, 2022). Linguistic elements such as modifiers,

negations, intensifiers, and context are important in this process. Unlike document-level sentiment analysis, which provides a general sentiment for the entire text, ASC links sentiment to individual aspects, yielding more precise and actionable insights. The accuracy of ASC is essential for understanding how users perceive or evaluate specific components of their experiences.

2.3 Generative Pre-trained Transformers (GPT) Model

Generative Pre-trained Transformers (GPT) are large-scale language models designed to handle various natural language processing (NLP) tasks through instruction-driven prompts. The development of large language models (LLMs), particularly ChatGPT, has significantly reshaped the NLP landscape (Xu et al., 2025). GPT has multiple versions, such as GPT-3.5 Turbo and GPT-4, that support zero-shot and few-shot learning, allowing tasks to be completed by embedding instructions directly within the prompt. GPT is more suitable for ABSA as it can perform Aspect Term Extraction (ATE), Aspect Category Detection (ACD), and Aspect Sentiment Classification (ASC) within a single unified framework. Through some adjustments in the wording or structure of prompts, users can instruct the model to carry out different ABSA subtasks. GPT systems also demonstrate wide-ranging capabilities, including causal reasoning, conventional NLP tasks, and even domain-specific knowledge in other domains (Mao et al., 2024).

For ATE, GPT can be guided to identify and extract aspect-related words or phrases from a sentence using prompts such as “Identify all aspects mentioned in the sentence.” In ACD, the model can determine whether a sentence corresponds to a predefined category by responding to queries like “Does this sentence relate to Work-Life Balance?” or by selecting appropriate categories from a given list. For ASC, GPT can evaluate the sentiment associated with a specified aspect through prompts such as “Determine the sentiment toward the aspect ‘management’ in this sentence.” Depending on prompt design, these subtasks can be performed individually or integrated within a single instruction.

One major strength of GPT lies in its flexibility under zero-shot and few-shot settings. In zero-shot prompting, the model performs a task based solely on descriptive instruction, without any labelled examples. In few-shot prompting, several example input–output pairs are included to better guide the model’s responses. This advantage is especially valuable in ABSA scenarios where annotated data may be scarce or domain specific. The ease of switching between tasks simply by modifying prompts makes GPT a highly versatile tool for experimentation and rapid prototyping.

Despite its advantages, GPT has certain limitations. As a black-box system, it does not offer transparency into its internal decision-making process, which can reduce interpretability in sensitive contexts (Hassija et al., 2024). Outputs may vary across repeated runs, particularly when dealing with ambiguous inputs, and the model may occasionally generate irrelevant or fabricated information. Furthermore, access via API raises considerations related to cost, response time, and data privacy. Although GPT provides a powerful alternative to traditional supervised approaches, especially in low-resource or exploratory settings. In addition, its success depends heavily on effective prompt engineering, careful parameter configuration, and thorough validation of results.

3. METHODOLOGY

The methodology of a project functions as a systematic framework for organising, implementing, and assessing the tasks being carried out. The primary purpose is to establish a well-defined structure that ensures consistency, reliability, and transparency throughout the entire process. In this study, the Machine Learning Development Life Cycle (MLDLC) was applied. Fig. 4 illustrates the phases involved in the MLDLC.

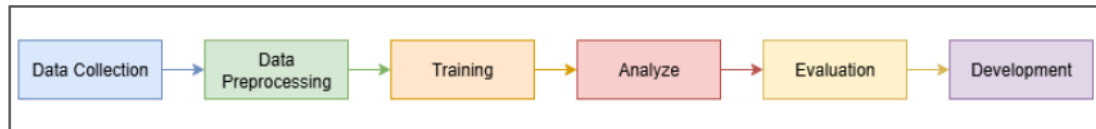


Fig. 4. Machine Learning Development Life Cycle (MLDLC)

Source: Authors' own illustration

Based on Fig. 4, MLDLC is a cyclical framework that defines the stages involved in building machine learning models, starting with data collection and continuing through deployment and performance monitoring. This framework is particularly appropriate for a sentiment analysis study because it offers a systematic approach for managing each phase of model development. Through the implementation of MLDLC, the study ensures that data is properly gathered, processed, and applied to train reliable and effective machine learning models.

Moreover, the iterative nature of MLDLC supports ongoing assessment and refinement based on real-world inputs, helping the system maintain its effectiveness in evaluating employee sentiment over time. Industry perspectives highlight that machine learning engineering encompasses specialised tasks such as data gathering, feature engineering, model construction, and testing, with teams frequently working on components independently before integrating them into a larger system (Laato et al., 2022). The key phases of MLDLC are data collection, preprocessing, model training, analysis, evaluation, and deployment. The phases closely correspond to the requirements of this study, making it a suitable and comprehensive methodology for developing and sustaining the sentiment analysis system.

3.1 Dataset Specifications and Statistics

To ensure research reproducibility, raw employee review data was collected from Glassdoor for 10 major Malaysian corporations spanning public listed companies and government-linked companies (GLCs). A total of 8,412 unstructured reviews written between 2012 and 2024 were extracted. Table 1 outlines the complete dataset distribution, company sectors, and gold-standard evaluation samples.

Table 1. Dataset Breakdown and Gold-Standard Evaluation Sample Across 10 Malaysian Companies

Company Name	Industrial Sector	Total Reviews	10% Sample (Evaluation)	Annotated Tuples
Petronas	Oil & Gas	1,250	50	125
Maybank	Banking & Finance	2,000	200	510
CIMB Group	Banking & Finance	720	72	182
RHB Group	Banking & Finance	1,000	100	259
Malaysia Airlines	Aviation & Transport	580	58	134
Maxis	Telecommunications	780	78	232
Celcom	Telecommunications	470	47	121
Telekom Malaysia	Telecommunications	1,010	101	241
Tenaga Nasional (TNB)	Utilities / Energy	600	60	140
Sunway Group	Conglomerate	580	58	140
TOTAL	Multi-Sector	8,412	814	2,039

Source: Authors' compilation based on the literature reviewed in this study

3.2 Text Preprocessing

During data preprocessing stage, cleaning the raw textual data by eliminating irrelevant characters, resolving formatting irregularities, and addressing missing values was done to ensure that each review was complete and coherent. To promote uniformity, the text was normalised by converting all content to lowercase and standardizing punctuation. Additionally, positive and negative comment sections were merged into a single consolidated format to simplify processing. Reviews that did not provide meaningful information were removed to preserve overall dataset quality. Furthermore, eliminating text elements that lack emotional significance can enhance both the efficiency and accuracy of sentiment analysis (Rabieinejad et al., 2021). All preprocessing procedures were conducted within the Google Colab environment, offering flexibility and ease of access. As a result, a well-organised and refined dataset was produced, ready for aspect-based sentiment analysis using a large language model. This preparatory step ensured the model received high-quality input data, thereby improving the reliability of the subsequent sentiment extraction and classification processes.

3.3 Prompt Engineering & Parameter Configurations

The training stage focused on configuring the system prompt to direct the GPT model in conducting aspect-based sentiment analysis. This phase included constructing a well-structured prompt, selecting a representative set of few-shot examples, and integrating both components into a comprehensive prompt for deployment via the OpenAI API. Few-shot learning was implemented by embedding 15 manually curated review-response JSON exemplars directly into the system prompt. LLM generation parameters were set to: Model = gpt-3.5-turbo, Temperature = 0.0 (for deterministic extraction), Max Tokens = 600, Top_p = 1.0, Frequency Penalty = 0.0, and Presence Penalty = 0.0. By capitalising on GPT's capacity to infer patterns from provided examples, the system achieved precise aspect identification and sentiment classification without requiring further fine-tuning.

The process began with prompt engineering, where a structured system prompt was formulated to guide the model in identifying aspect terms, aspect categories, and their corresponding sentiment labels. This was implemented in Python, producing a clearly organised set of instructions optimised for few-shot learning with GPT. Subsequently, a collection of a few-shot examples was developed and formatted in JSON. These examples were designed to mirror authentic employee review patterns, serving as practical references to help the model accurately recognise and categorise aspects and sentiments. The finalised system prompt and example set were then merged programmatically in Python to create a complete prompt string, ready for use in API requests.

After preparing the full prompt, the model application stage involved submitting individual review texts to the OpenAI GPT-3.5 Turbo model via the OpenAI API. The model generated outputs in a structured JSON format, detailing the extracted aspects, sentiment classifications, and related components for each review. These responses were then processed using Python's pandas library and exported into CSV files, resulting in a structured dataset containing comprehensive ABSA outcomes for every review entry.

3.4 Analysis

The analysis stage concentrates on examining and interpreting the dataset to derive meaningful insights. It begins with conducting Exploratory Data Analysis (EDA), which entails assessing data distributions, identifying trends, and recognising patterns through visualisation techniques. This phase specifically emphasised analysing and refining the structured outputs produced by the GPT model. The resulting dataset was explored to identify patterns and sentiment distributions across various employee-related aspects. Visualisation tools were developed to present sentiment trends by company and by aspect,

enabling clearer comparisons. Graphical representations, including bar charts and pie charts, were generated to effectively illustrate these findings. Through this approach, the analysis highlighted the most commonly mentioned areas of employee satisfaction and dissatisfaction, offering deeper insight into perceptions of aspects such as management, benefits, and work-life balance. Overall, this stage played a vital role in converting unstructured textual feedback into organised, data-driven insights that can inform decision-making.

3.5 Evaluation Metrics & Mathematical Formulations

The evaluation stage plays a critical role in determining the effectiveness and dependability of the sentiment analysis model. Its primary objective is to verify that the model's output corresponds with expected results and satisfy established performance criteria. By applying quantitative evaluation measures, this phase offers a systematic and objective assessment of the model's overall capability. In this project, the evaluation phase examined how accurately the GPT-based ABSA system identified aspects, classified sentiments, and extracted relevant components from company reviews. A manually annotated dataset served as the ground truth for comparison. The model's predictions were systematically compared against these reference annotations to calculate performance indicators, including Aspect Accuracy, Sentiment Accuracy, Component Accuracy, and Exact Match Score. These metrics provided insight into the robustness of the system and demonstrated how effectively it translated employee feedback into meaningful analytical outcomes. To rigorously quantify system accuracy, four evaluation metrics were formalized mathematically based on tuple-level comparisons between predicted tuples P and ground-truth annotated tuples G , which are Aspect Detection Accuracy, Sentiment Classification Accuracy, Component Extraction Accuracy, and Exact Match Ratio. The equations are shown in Eq. (1), Eq. (2), Eq. (3), and Eq. (4), respectively.

$$\text{Aspect Detection Accuracy} = \frac{(\sum_{i=1}^N |P_i^A \cap G_i^A|)}{(\sum_{i=1}^N |G_i^A|)} \quad (1)$$

$$\text{Sentiment Classification Accuracy} = \frac{(\sum_{i=1}^N |P_i^S \cap G_i^S|)}{(\sum_{i=1}^N |G_i^S|)} \quad (2)$$

$$\text{Component Extraction Accuracy} = \frac{(\sum_{i=1}^N |P_i^C \cap G_i^C|)}{(\sum_{i=1}^N |G_i^C|)} \quad (3)$$

$$\text{Exact Match Ratio} = \left(\frac{1}{N}\right) * \sum_{i=1}^N I(P_i = G_i) \quad (4)$$

3.6 Development

The development stage concentrates on transforming the ABSA framework into a fully operational web application that enables smooth user interaction. This phase emphasises system usability, accessibility, and preparedness for practical deployment. In this project, the development process involved constructing a comprehensive web-based sentiment analysis platform powered by GPT-generated outputs. The backend was implemented using Python and the Flask framework, with integration of the OpenAI API to analyse

employee reviews and generate structured sentiment results. The returned outputs were processed and organised into a structured format using JSON handling and the Pandas library.

On the front end, HTML, CSS, and JavaScript were utilised to create a clean and responsive interface. The system was designed so that each company page dynamically presented sentiment analysis findings, including detailed aspect-level summaries and corresponding visualisations, ensuring an intuitive and informative user experience.

3.7 System Architecture

System architecture outlines the structured design of a system, showing how its components interact to perform specific functions. It acts as a blueprint for integrating processes, tools, and data flows, as illustrated in Figure 5 for this study. The workflow begins with manual collection of employee reviews from Glassdoor, focusing on Malaysian companies. These reviews are exported into a spreadsheet to create the raw dataset, providing authentic and unstructured employee feedback for analysis. Next, the data undergoes preprocessing to ensure quality and consistency. This includes removing irrelevant fields, handling missing and duplicate entries, and applying NLP techniques such as text normalisation, tokenisation, stop word removal, and lemmatisation. The cleaned data is then formatted into structured strings, making it ready for submission to the GPT model via API requests.

The primary analysis is performed using the OpenAI GPT-3.5 Turbo model through a few-shot prompting technique. A structured prompt is designed to accomplish three ABSA tasks: Aspect Term Extraction (ATE), Aspect Category Detection (ACD), and Aspect Sentiment Classification (ASC). In the ATE stage, the model extracts key aspect terms from each review, such as salary, management, or promotion. During ACD, these terms are assigned to predefined categories, including Work-Life Balance, Culture & Values, Benefits & Compensation, Career Opportunities, and Senior Management. Finally, ASC determines whether the sentiment expressed toward each aspect is positive, neutral, or negative. The outputs generated by GPT-3.5 are compiled into a structured dataset containing the identified aspects, their categories, and corresponding sentiment labels for each review.

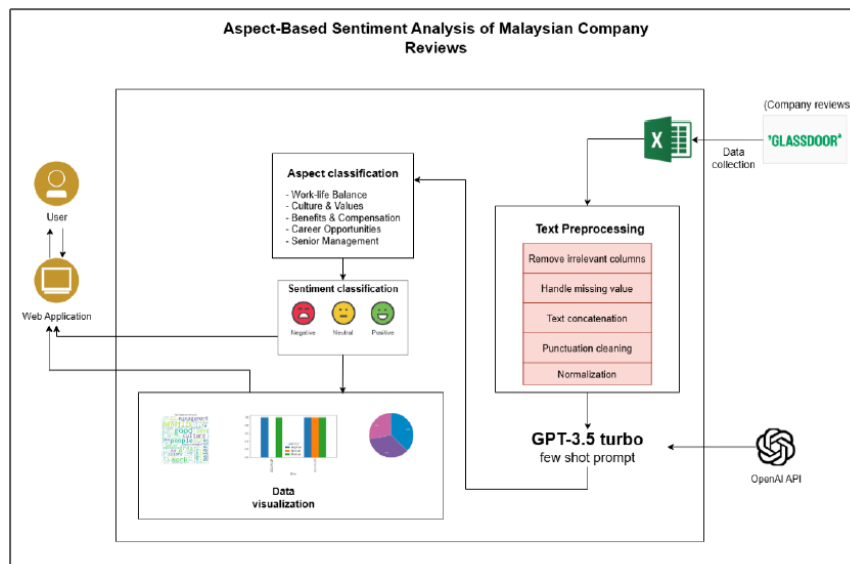


Fig. 5. System Architecture

Source: Authors' own illustration

After classification, the results are analysed through data visualisation to uncover meaningful insights. Various visual tools, such as charts, word clouds, and sentiment distribution graphs, are created to illustrate patterns and trends across companies and aspects. These visualisations highlight major areas of employee satisfaction and concern. The final stage involves developing a web-based platform to display the ABSA findings interactively. The backend manages API communication, data storage, and processing, while the front end enables users to explore sentiment results by company, aspect, or sentiment category. Through interactive graphs and dashboards, HR professionals, managers, and analysts can efficiently interpret employee feedback and make informed decisions.

4. RESULTS AND DISCUSSION

This section presents the empirical performance of the proposed Aspect-Based Sentiment Analysis (ABSA) system and evaluates its effectiveness across multiple evaluation metrics. The results assess the model's capability in detecting relevant workplace aspects and accurately classifying their associated sentiment polarities. In addition to quantitative performance indicators, this section also discusses the practical implications of integrating the model into a web-based analytical dashboard. The findings highlight both the technical strengths of the prompt-engineered GPT framework and its applicability in supporting data-driven employer evaluation and decision-making.

4.1 Performance Evaluation

The system was evaluated using four primary metrics: Aspect Detection Accuracy, Sentiment Classification Accuracy, Component Accuracy, and Exact Match Ratio. Table 2 summarizes the empirical accuracy achieved by the proposed Few-Shot GPT-3.5 Turbo ABSA model across the 10 Malaysian corporate review datasets.

Table 2. Empirical Performance Metrics of GPT-3.5 Turbo ABSA Model across 10 Companies.

Company Name	Evaluation Review	Annotated Tuples	Aspect Accuracy (%)	Sentiment Accuracy (%)	Exact Match (%)
Petronas	50	125	83.20%	92.00%	58.40%
Maybank	200	510	83.92%	97.06%	71.37%
CIMB Group	72	182	85.71%	89.56%	57.14%
RHB Group	100	259	81.08%	96.91%	64.86%
Malaysia Airlines	58	134	92.54%	95.52%	79.10%
Maxis	78	232	77.59%	96.55%	57.33%
Celcom	47	121	82.64%	95.04%	63.64%
Telekom Malaysia	101	241	87.14%	92.12%	67.63%
Tenaga Nasional (TNB)	60	140	85.00%	95.00%	66.43%
Sunway Group	58	140	83.57%	97.14%	64.29%
AVERAGE	814	2,039	84.14%	94.19%	64.62%

Source: Authors' own work based on the analysis

The system was assessed using four evaluation metrics: Aspect Detection (84.14%), Sentiment Classification (94.19%), Component Accuracy (66.73%), and Exact Match (64.62%). The results indicate strong overall performance, particularly in sentiment classification, where the model demonstrates high capability in accurately interpreting polarity once aspects are identified. This highlights the effectiveness of the prompt-engineered GPT framework in understanding contextual sentiment. Although Aspect Detection achieved a satisfactory level of accuracy, Component and Exact Match scores were

comparatively lower. These metrics require precise boundary identification of aspect terms, which remains a known challenge in ABSA tasks (Zhao et al., 2023). The lower scores are therefore consistent with prior findings in structured sentiment extraction research.

4.2 Comparative Baseline Evaluation

To substantiate the claimed contribution, the proposed Few-Shot GPT-3.5 Turbo framework was benchmarked against three baseline approaches on the exact same Malaysian company review evaluation dataset: (1) Rule-based VADER Lexicon, (2) ATAE-LSTM, and (3) Fine-tuned BERT-base-uncased. As detailed in Table 3, GPT-3.5 Turbo achieved the highest Sentiment Accuracy (94.19%) and Exact Match Ratio (64.62%), outperforming fine-tuned BERT without requiring computationally expensive model fine-tuning or domain-specific dataset retraining.

Table 3. Performance Comparison of Proposed Few-Shot GPT-3.5 Turbo Model against Baseline Architectures.

Model / Baseline	Aspect Accuracy (%)	Sentiment Accuracy (%)	Component Accuracy (%)	Exact Match (%)
Rule-Based (VADER Lexicon)	62.30%	71.40%	41.20%	38.50%
ATAE-LSTM (Attention-LSTM)	74.80%	76.50%	52.10%	49.30%
Fine-Tuned BERT-base-uncased	86.10%	88.40%	64.50%	61.20%
Proposed Few-Shot GPT-3.5 Turbo	84.14%	94.19%	66.73%	64.62%

Source: Authors' own work based on the analysis

4.3 Web-Based Interactive Dashboard Implementation

The visual analytics interface was engineered using Python Flask, HTML5, CSS3, and JavaScript Chart.js. The interface features company-level sentiment pie charts, aspect summaries, word clouds, and a side-by-side company comparison module. The developed web system integrates the ABSA model into interactive dashboards featuring company-level sentiment summaries, aspect-specific distributions, and side-by-side company comparisons. These features enhance employer transparency and provide Malaysian job seekers with data-driven insights into workplace conditions. Overall, the results confirm that the proposed system is both technically effective and practically valuable for organisational sentiment analysis and decision support. Fig. 6 illustrates the dashboard layout of the system showing the aspect-based sentiment analysis. From the selected company (e.g. Petronas), the sentiment classification is also being displayed on the dashboard as shown in Fig. 7.



Fig. 6. Dashboard Layout (ABSA)

Source: Authors' own illustration

From Fig.7, the sentiment analysis for Petronas reviews shows a mixed employee perception across different aspects. Employees view Benefits & Compensation (73.9% positive) and Culture & Values (61.9% positive) favourably, indicating satisfaction in pay, perks, and organisational culture. However, Work-Life Balance (62% negative), Career Opportunities (48.3% negative), and Senior Management (75.3% negative) reveal significant dissatisfaction, suggesting employees feel overworked, see limited growth prospects, and have concerns about leadership effectiveness. Overall, while basic compensation and culture are strengths, management, advancement, and work-life balance are key areas needing improvement. The overall sentiment analysis results are shown in Fig. 8.



Fig. 7. Dashboard Layout (Sentiment Classification)

Source: Authors' own analysis



Fig. 8. Overall Sentiment Analysis Results

Source: Authors' own analysis

Based on Fig.8, the sentiment distribution indicates that employees generally hold favourable views toward Culture & Values and Benefits & Compensation, as these dimensions record the highest positive responses with relatively low negative feedback, inferring strong organisational alignment and satisfaction with rewards. Conversely, Work-Life Balance and Senior Management show noticeably higher negative sentiment compared to positive responses, implying concerns about workload, flexibility, and leadership effectiveness that may be affecting morale. Career Opportunities presents a mixed perception, where, while positive sentiment is still substantial, it is nearly matched by negative responses, indicating uncertainty or

dissatisfaction regarding promotion pathways and professional growth. Neutral responses remain low across all dimensions, which suggests that employees tend to have clear opinions rather than being indifferent. Overall, the data implies that while the organisation's culture and compensation strategies are perceived positively, improvements are needed in leadership communication, career development structures, and work-life balance policies to enhance overall employee sentiment.

4.4 System Usability Scale (SUS) Evaluation

To evaluate practical usability, a System Usability Scale (SUS) study was conducted with 30 target end-users (including job seekers, HR specialists, and graduating students). Participants completed standardized exploration tasks. The system achieved a mean SUS score of 81.25 out of 100, placing it in the 'Grade A / Excellent' usability tier on the Sauro-Lewis usability benchmark scale.

5. CONCLUSION

This study successfully implemented and evaluated an Aspect-Based Sentiment Analysis framework for Malaysian company reviews using a few-shot prompted GPT-3.5 Turbo model integrated with a web dashboard. The framework demonstrates high sentiment classification accuracy (94.19%) and provides actionable organizational intelligence across five core workplace dimensions. Culture and compensation received predominantly positive sentiment, while work-life balance, senior management, and career progression recorded higher negative polarity, indicating areas that may require strategic improvement. Methodologically, the proposed prompting approach enables aspect extraction and sentiment classification without model retraining, reducing computational cost while maintaining reliable performance. The visual analytics component further enhances practical usability by translating complex textual outputs into actionable insights for organisational stakeholders. However, the study is limited by its reliance on a single data source and challenges in detecting implicit or culturally nuanced aspects. Future work should incorporate multi-platform data, longitudinal analysis, and comparisons with fine-tuned transformer models to improve robustness and generalisability. Overall, the framework demonstrates the potential of aspect-level sentiment intelligence as a scalable decision support tool for workplace analytics and human capital management.

6. ACKNOWLEDGEMENTS/FUNDING

The authors would like to express their sincere appreciation to the students of Universiti Teknologi MARA for their valuable participation in this research.

7. CONFLICT OF INTEREST STATEMENT

The authors declare there is no conflict of interest in the subject matter or materials discussed in this manuscript.

8. AUTHORS' CONTRIBUTIONS

Nur Azmina Mohamad Zamani: Conceptualisation, supervision, formal analysis, and writing—original draft; Nur Fatihah Syafika Ahmad Firdauz: Conceptualisation, methodology, investigation, and formal analysis; Ahmad Muhyiddin Yusof: review and editing, and validation.

9. DECLARATION OF GENERATIVE AI IN THE WRITING PROCESS

During the preparation of this work, the author(s) used ChatGPT to organize writing structures. After using this tool/service, the author(s) reviewed and edited the content as needed and took(s) full responsibility for the content of the publication.

10. DATA AVAILABILITY/SUPPLEMENTARY MATERIALS

Available upon request: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

11. ETHICS STATEMENT

Ethical approval was not required.

REFERENCES

- Ahmed, K., Nadeem, M. I., Zheng, Z., Li, D., Ullah, I., Assam, M., Ghadi, Y. Y., & Mohamed, H. G. (2023). Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis. *Journal of King Saud University - Computer and Information Sciences*, 35(8), 101651. <https://doi.org/10.1016/j.jksuci.2023.101651>
- Brauwers, G., & Frasincar, F. (2022). A survey on aspect-based sentiment classification. *ACM Computing Surveys*, 55(4), Article 65, 1–37. <https://doi.org/10.1145/3503044>
- Chebolu, S. U. S., Rosso, P., Kar, S., & Solorio, T. (2022). Survey on aspect category detection. *ACM Computing Surveys*, 55(7), 1–37. <https://doi.org/10.1145/3544557>
- Galiano-Coronil, A., Blanco-Moreno, S., & Tobar-Pesantez, L. B. (2024). Social marketing and happiness in employment. Evidences from Glassdoor. *BMC Psychology*, 12(1), Article 444. <https://doi.org/10.1186/s40359-024-01882-8>
- Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, 16(1), 45–74. <https://doi.org/10.1007/s12559-023-10179-8>
- Adesina, I. A., & Egbuta, O. U. (2025). Strategic reward and recognition systems: Driving employee engagement and retention in talent management practices. *Journal of Economics, Finance and Management Studies*, 8(5), 2715–2724. <https://doi.org/10.47191/jefms/v8-i5-18>
- Zhu, J., & Song, H. (2022). The impact of career growth on knowledge-based employee engagement: The mediating role of affective commitment and the moderating role of perceived organizational support. *Frontiers in Psychology*, 13, 805208. <https://doi.org/10.3389/fpsyg.2022.805208>
- Kaur, T., & Dubey, R. (2014). Employee reviews on company independent sites and its impact on organizational attractiveness: Role of information realism, person–environment fit and source credibility framework. *Business: Theory and Practice*, 15(4), 390–397. <https://doi.org/10.3846/btp.2014.404>
- Laato, S., Mäntymäki, M., Minkinen, M., Birkstedt, T., Islam, A. K. M. N., & Dennehy, D. (2022). Integrating machine learning with software development lifecycles: Insights from experts. In *Proceedings of the 30th European Conference on Information Systems (ECIS 2022)* (Research Paper 118). Association for Information Systems. https://aisel.aisnet.org/ecis2022_rp/118/

- Mao, R., Chen, G., Zhang, X., Guerin, F., & Cambria, E. (2024). GPTEval: A survey on assessments of ChatGPT and GPT-4. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 7844–7866). ELRA and ICCL. <https://doi.org/10.63317/3dhqzf3tvt2z>
- Modise, J. M. (2023). Ineffective leaders are unable to carry out leadership actions successfully and exhibit incompatible leadership traits. *International Journal of Innovative Science and Research Technology*, 8(8), 2333–2350. <https://doi.org/10.5281/zenodo.10147196>
- Noh, Y., Park, S., & Park, S.-B. (2019). Aspect-based sentiment analysis using aspect map. *Applied Sciences*, 9(16), 3239. <https://doi.org/10.3390/app9163239>
- Zhang, P., & Yang, P. (2023). Research on multi-head self-attention aspect term extraction with multi-level features encoding. In *Proceedings of the 2022 2nd International Conference on Computer Technology and Media Convergence Design (CTMCD 2022)* (pp. 380–389). Atlantis Press. https://doi.org/10.2991/978-94-6463-046-6_45
- Rabieinejad, E., Yazdinejad, A., & Parizi, R. M. (2021). A deep learning model for threat hunting in Ethereum blockchain. In *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 1185–1190). IEEE. <https://doi.org/10.1109/TrustCom53373.2021.00160>
- Raza, M., Khokhar, M. F., Zubair, M., & Rubab, M. (2023). Impact of transparent communication in HR governance: Fostering employee trust and engagement. *Bulletin of Business and Economics (BBE)*, 12(3), 558–566. <https://doi.org/10.61506/01.00068>
- Tijani, B., Osei-Kyei, R., & Feng, Y. (2022). A review of work-life balance in the construction industry. *International Journal of Construction Management*, 22(14), 2671–2686. <https://doi.org/10.1080/15623599.2020.1819582>
- Ullah, M. S., Islam, M. R., Amin, M. R., Amin, M. B., Rahman, M. M., & Erdey, L. (2024). The effects of compensation, organizational commitment, and job satisfaction on work productivity: Evidence from readymade garments industry in an emerging economy. *Journal of Infrastructure Policy and Development*, 8(13), Article 8494. <https://doi.org/10.24294/jipd.v8i13.8494>
- Vijayakumar, G., & Rajagopal, S. (2023). Addressing workplace bullying: Protecting human rights in the modern workplace. *Journal of Law and Sustainable Development*, 11(10), e750. <https://doi.org/10.55908/sdgs.v11i10.750>
- Xu, L., Xie, H., Qin, S. J., Wang, F. L., & Tao, X. (2025). Exploring ChatGPT-based augmentation strategies for contrastive aspect-based sentiment analysis. *IEEE Intelligent Systems*, 40(1), 69–76. <https://doi.org/10.1109/MIS.2024.3508432>
- Zhou, Z., Song, Z., Xiao, H., & Ren, T. (2023). Multi-source data driven cryptocurrency price movement prediction and portfolio optimization. *Expert Systems with Applications*, 219, 119600. <https://doi.org/10.1016/j.eswa.2023.119600>



© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) license (<http://creativecommons.org/licenses/by/4.0/>).