



# Customer Perspectives on Hotel Reviews in DHAKA

## **Supervised By**

**Dr. Imran Mahmud**

Head & Associate Professor  
Department of Software Engineering  
Daffodil International University

## **Submitted By**

**Omar Faruk Himel**

**Id: 221-35-846**

Department of Software Engineering  
Daffodil International University

This thesis report has been submitted in fulfilment of the requirements for the Degree of Bachelor of Science in Software Engineering.

## APPROVAL

This thesis titled on “Customer Perspectives on Hotel reviews in Dhaka”, submitted by **Omar Faruk Himel (ID: 221-35-846)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

### BOARD OF EXAMINERS



---

**Dr. Imran Mahmud**  
**Professor & Head**  
Department of Software Engineering  
Faculty of Science and Information Technology Daffodil International University

**Chairman**



---

**Afsana Begum**  
**Assistant Professor**  
Department of Software Engineering  
Faculty of Science and Information Technology  
Daffodil International University

**Internal Examiner 1**



---

**Md. Shohel Arman**  
**Assistant Professor**  
Department of Software Engineering  
Faculty of Science and Information Technology  
Daffodil International University


**Internal Examiner 2**



---

**Nadira Islam**  
**Assistant Professor**  
Department of Software Engineering  
Faculty of Science and Information Technology  
Daffodil International University

**Internal Examiner 3**



---

**Md Manowarul Islam**  
**Professor**  
Department of Computer Science and Engineering  
Jagannath University, Bangladesh

**External Examiner**

# Customer Perspectives on Hotel Reviews in DHAKA

Omar Faruk Himel

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

## DAFFODIL INTERNATIONAL UNIVERSITY

### DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name : Omar Faruk Himel  
Date of Birth : 01-01-2002  
Title : Customer Perspectives on Hotel Reviews in DHAKA  
Academic Session : 2022-2025

I declare that this thesis is classified as:

- CONFIDENTIAL (Contains confidential information under the Official Secret Act 1997)\*
- RESTRICTED (Contains restricted information as specified by the organization where research was done)\*
- OPEN ACCESS I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:

*Omar Faruk Himel*

\_\_\_\_\_  
(Student's Signature)

221-35-846

\_\_\_\_\_  
Student ID  
Date: 27/ November /2025

*Dr. Imran Mahmud*

\_\_\_\_\_  
(Supervisor's Signature)

Dr. Imran Mahmud

\_\_\_\_\_  
Name of Supervisor  
Date: 27/ November /2025

NOTE : \* If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.



## SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science.

A handwritten signature in black ink, appearing to be "Imran Mahmud", is written above a horizontal line.

(Supervisor's Signature)

Full Name : Dr. Imran Mahmud

Position : Head & Associate Professor, Department of Software Engineering.

Date : 27/ November /2025



## STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

*Omar Faruk.*

---

(Student's Signature)

Full Name : OMAR FARUK HIMEL

ID Number : 221-35-846

Date : 27 November 2025

# Customer Perspectives on Hotel Reviews in DHAKA

Omar Faruk Himel

Thesis submitted in fulfillment of the requirements  
for the award of the degree of  
Bachelor of Science

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

November 2025

## ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Almighty Allah who gave me the power, patience, and wisdom to complete this research work successfully.

I would like to give my heartfelt thanks to my supervisor whose support, motivation, and assistance were a great help. advice have played a significant role throughout the completion of this thesis. Their ongoing help and useful feedback have helped me to improve my ideas and achieve the

goals of this research.

I would like to thank all the members of the faculty of the Department of Software Engineering, Daffodil International University, through their guidance, assistance, and the academic. experience they have provided me with during my undergraduate life.

I would like to use this opportunity to thank my friends and co-workers in research whose. inspiration, technical support, and thought-provoking discussions helped me to address challenges during this project.

To sum up, I would like to express my sincere gratitude to my parents and my family. because of their undying love, encouragements, and prayers. Their uncompromising trust in me has been the driving force to all my achievements.

## **DEDICATION**

To change the overall look of your document, choose new theme elements on the Page Layout tab. To change the looks available in the Quick Styles gallery, use the Change Current Quick Style Set command. Both the Themes gallery

## **ABSTRACT**

Customer sentiment based on hotel reviews is crucial to the quality improvement process in the digital hospitality context, but informal and multilingual feedback is difficult to analyze. This paper deals with the issue of correctly identifying positive and negative guest reviews based on real life hotel reviews in Dhaka, which was gathered through a Google API scraper in order to obtain quantity and authenticity. The suggested solution is a hybrid one with the use of Long Short-Term Memory (LSTM) neural networks and Term Frequency-Inverse Document Frequency (TF IDF) features. The methodology includes the comprehensive data cleaning, tokenization, feature engineering, and a dual-branch neural network: sequential LSTM modeling of contextual cues and TF IDF encoding of global lexical detail. The model achieves a high classification accuracy of about 95 percent and the precision and recall are balanced between the two levels of sentiment. The findings validate the usefulness of sequence-based deep learning with sparse lexical features in robust sentiment interpretation of diverse, automatically-collected hotel review datasets.

## TABLE OF CONTENT

<b>DECLARATION</b>	
<b>TITLE PAGE</b>	
<b>ACKNOWLEDGEMENTS</b>	<b>viii</b>
<b>Dedication</b>	<b>ix</b>
<b>ABSTRACT</b>	<b>x</b>
<b>TABLE OF CONTENT</b>	<b>xi</b>
<b>LIST OF FIGURES</b>	<b>xiv</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
1.1 Problem Identification	1
1.2 Aim and Objective	1
1.3 Research Scope	2
1.4 Research Contribution	3
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>5</b>
2.1 Concept of Customer Perception in the Hospitality Industry	5
2.2 Sentiment Analysis in the Context of Hotel Reviews	6
2.3 Research on Online Customer Reviews in the Hotel Sector	7
2.4 Existing Models and Analytical Approaches	9
2.4.1 Machine Learning Approaches	9
2.4.2 Deep Learning and Transformer-Based Approaches	10
2.4.3 Hybrid and Explainable AI (XAI) Methods	11
2.5 Identified Research Gaps	12
2.5.1 Gaps in Dhaka-Specific Hotel Review Research	12
2.5.2 Handling Linguistic and Cultural Diversity	12
2.5.3 Challenges in Contextual Understanding (Sarcasm, Irony)	13

2.5.4	Dataset Limitations and Imbalance Issues	13
2.5.5	Need for Better Interpretability and Adaptive Models	14
2.6	Chapter Summary	14
<b>CHAPTER 3 METHODOLOGY</b>		<b>16</b>
3.1	Proposed Methodology	17
3.2	Data Collection	18
3.2.1	Data Labelling	20
3.2.2	Pre-processing Steps	21
3.3	Data Splitting	23
3.4	Model Architecture	25
3.5	Model Training	28
3.6	Evaluation	30
3.7	Chapter Summary	32
<b>CHAPTER 4 RESULTS AND DISCUSSION</b>		<b>33</b>
4.1	Performance Metrics	34
4.1.1	Classification Report	35
4.1.2	Confusion Matrix	36
4.1.3	Experimental Result Analysis	37
4.2	Accuracy and Loss Plots	38
4.3	Chapter Summary	39
<b>CHAPTER 5 CONCLUSION</b>		<b>41</b>
5.1	Findings of the Study	41
5.2	Key Contributions	42
5.3	Future Works	43



## LIST OF FIGURES

Figure 1.1 Research Scope	2
Figure 3.1: Methodology Diagram	18
Figure 3.2 : word cloud based only on the negative reviews	20
Figure 3.3: Data Distribution of the Dataset	21
Figure 3.4: Model Evaluation Results	31
Figure 4.1: classification report	36
Figure 4.2: Confusion Matrix	37
Figure 4.3: Accuracy and Loss Plots	39

# CHAPTER 1

## INTRODUCTION

Hotel business is undergoing a digital revolution at a very rapid pace and online reviews are becoming one of the biggest factors that affect consumer choice and service enhancement strategies. To achieve scale and authenticity, automated data collection of hotel reviews was used in this research via a Google API scraper, which allowed the researcher to analyze a large and representative sample of customer reviews. Users are becoming more and more dependent on online platforms to post and review hotel experiences, which are influencing reputations and reservations worldwide. In such a data-driven world, it is essential to derive practical insights out of unstructured online reviews to enable hotel managers who seek to improve on service quality. The recent research has emphasized the necessity of studying customer-generated content to learn the factors of satisfaction, weaknesses of the services, and changing expectations of the guests.

### 1.1 Problem Identification

Although it has made great strides in sentiment analysis, a lot of the current methods are not effective in the classification of customer sentiment in real-world, multilingual hotel reviews data. The traditional machine learning, rule-based approaches, or the analysis of individual aspects of reviews have been used in the past, which can be ineffective in considering the context and even the nuances of emotions. There are still gaps in finding a balanced performance in classification and extraction of insights that will be consistent with both operational requirements and dynamic guest language.

### 1.2 Aim and Objective

The proposed project aims to overcome these shortcomings by creating a superior sentiment analysis framework that would be used to classify hotel reviews as binary (positive/negative) in the hospitality industry of Dhaka. Based on the data obtained

automatically through the use of Google review websites, and applying deep learning (LSTM) with statistical text vectors (TF-IDF), the study will enhance classification accuracy and offer a deeper and more data-driven insight into guest experiences. The result assists the managers and policymakers of the hotel in tracking the satisfaction, in filling service gaps, and in maintaining competitive advantages in an online market.

### 1.3 Research Scope

The project includes the end-to-end analysis of hotel review data, including the acquisition (through Google API scraper) and pre-processing of user-generated content to model development, assessment, and sentiment results interpretation. The binary sentiment levels (positive/negative) are only investigated, and the methodological innovation, performance evaluation, and the practical implications to the local hotel management are considered. The strategy is aimed at being generalized to other similar datasets and settings.

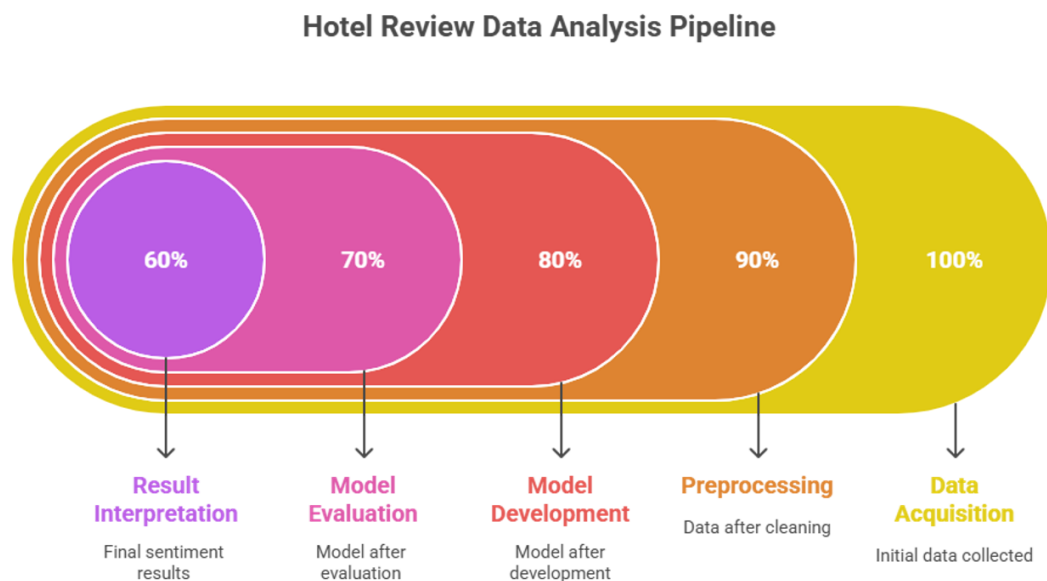


Figure 1.1 Research Scope

- **Data Acquisition (100%):** This is the first step where raw hotel review data is acquired, which in your case is done by an automated Google API scraper. It is important to have a wide and natural dataset to obtain credible analysis outcomes.

- **Preprocessing (90%):** This stage involves cleaning of the gathered data to eliminate errors, duplicates, irrelevant data and standardization of text format. Quality preprocessing improves the quality of data and prepares it to model the sentiment with precision.
- **Model Development (80%):** It involves the construction of LSTM-TF-IDF hybrid sentiment analysis model, using the cleaned data. The step develops both the engineering of features and the architecture of the neural network to facilitate a strong learning.
- **Model Evaluation (70%):** This is a step where the model that has been trained is tested in terms of accuracy and reliability with separate validation data or test data. Precision, recall, and F1-score are the main measurements that are conducted to guarantee the model is achieving the desired performance.
- **Result Interpretation (60%):** This is the last step, which involves interpretation and presentation of the model results- interpretation of the sentiment results and the discussion of the practical implications of the findings to the hotel management and decision-makers. Such results can be used to improve operations and customer satisfaction.

#### 1.4 Research Contribution

- **Automated, Large-Scale Data Collection:** This thesis creates an effective pipeline that will obtain hotel review data on a large scale using automated Google API scraping. The automation will help the project to achieve authenticity and diversity that are required in the high-quality sentiment analysis.
- **Thorough Data Pre-Processing:** The paper proposes a strict text cleaning and pre-processing plan, which is specific to informal and multilingual and user-generated hotel reviews. The step improves the integrity and reliability of data as a model training basis.
- **Hybrid Deep Learning Framework:** A new hybrid sentiment classification model is suggested, that combines LSTM networks with sequential text comprehension

with TF-IDF with global term relevance. This hybrid is a combination of both the deep learning and the traditional vectorization in one architecture.

- **Binary Sentiment Focus:** The study fills a gap in current literature, as many studies tend to ignore the fact that binary sentiment classification (positive/negative) is simple and operationally significant in business applications in the hospitality industry.
- **Performance Assessment and Validation:** The thesis undertakes an extensive analysis on real-world hotel data, which reports balanced and high accuracy, precision and recall. This validation shows that the approach is practically effective and can be generalized to the situation in the Dhaka hotel.
- **Actionable Insights to Industry:** The study will be translated into actionable insights to the hotel management in Dhaka to promote evidence-based decision-making, improvement of service quality, and monitoring of customer satisfaction.
- **Reusable Methodological Framework:** The data pipeline and modeling strategy that was developed are reusable, which provides a generalizable template to sentiment analysis in similar datasets or industries with similar challenges.

The contributions contribute to the academic and practical knowledge of applied sentiment analysis in hospitality industry, which facilitates scalable, precise, and practical application of guest feedback to managerial and policy advantage. benefit.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Concept of Customer Perception in the Hospitality Industry

The customer perception in the hospitality industry can be defined as how the customers perceive and analyse their experiences with hotels, restaurants, or other service-oriented companies. It is influenced by various factors, including quality of services, customisation, employee conduct, hygiene, cost, facilities, as well as the general ambience that a guest is offered during his or her stay. The intangibility of services in the hospitality industry, with both production and consumption occurring in one place, underscores the significance of these perceptions, as customers can easily evaluate the overall experience based not just on the tangible products but also on the sense of being welcomed, taken care of, and appreciated by the employees and the management.

Service encounters, which include check-ins, staff interaction, or even resolution of problems are crucial in the creation of impressions. It has been proposed that pleasant experiences and unforgettable personalized experiences may result in positive perceptions, which translate into repeat visits and long-term loyalty. Simultaneously, a negative experience or comment can change the overall perception of a customer towards a property or a brand considerably. Meeting and managing expectations, maintaining consistency in quality, and actively responding to the comments of guests play a very important role in ensuring that the maximum satisfaction and the least amount of differences between the expected and perceived quality of service are achieved.

The digital age has increased the level of customer perception through the extensive application of online reviews and social media networks, and reputation management has become a necessity in hospitality businesses. Best hospitality management is the ability to know these perceptions, to constantly check on the feedbacks and to adjust the strategies to provide individualized experiences with emotional appeal that makes a business stand out in a competitive market.

## 2.2 Sentiment Analysis in the Context of Hotel Reviews

Sentiment analysis has become one of the fundamental analytical methods of studying customer perceptions expressed in online hotel reviews. As the digital platforms, including TripAdvisor, Booking.com, and Google Reviews, grow fast, customers now present their evaluations in detail, including their emotional reactions, satisfaction rates, and expectations of the hotel services. Sentiment analysis helps the researcher to transform this huge amount of unstructured information into structured information, thus enhancing the accuracy of the customer perception research in the hospitality sector [1], [2].

In the hotel research field, sentiment analysis is mostly concerned with finding out the polarity (positive, negative, or neutral), and finding out the sentiment on service specific attributes like cleanliness, staff attitude, food quality, room comfort, and pricing. The use of machine learning and deep learning models such as Support Vector Machines (SVM), LSTM networks, and transformer-based models has gained considerable popularity because of their capacity to identify linguistic patterns and contextual meaning in the words of customers [3], [4]. Specifically, aspect-based sentiment analysis allows identifying sentiments related to certain service attributes instead of the tone of the reviews, which gives hotels a more practical insight on what strengths and weaknesses they perform [1], [3].

Sentiment analysis in the hotel industry is a very important managerial tool. According to previous studies, the sentiment scores calculated based on online reviews are highly correlated with customer satisfaction ratings, booking intentions, and brand loyalty [4], [5]. Also, sentiment trends enable managers to pinpoint operational problems that are recurring, including cleanliness problems, service delays, or employee behavior problems, which allow managers to intervene in time and improve services continuously. In markets like Dhaka where the hotels are highly diverse in terms of quality and pricing and customers have diverse expectations, sentiment analysis is a very important source of evidence to determine the gaps in service and benchmark competitive performance [6].

In general, the inclusion of sentiment analysis in hotel review research increases the level of analysis of customer perception studies. It facilitates the transition towards the data-

driven management of hospitality, where customer-created reviews are not only feedback sources, but also strategic assets of service innovation and competitive distinction [2].

### **2.3 Research on Online Customer Reviews in the Hotel Sector**

Research of online customer reviews in the hotel industry have made online reviews an important type of electronic word-of-mouth (eWOM) that influences customer behaviour and hotel performance. Empirical evidence based on big data of ratings and textual comments demonstrates that review valence and content are strongly linked to revenue per available room (RevPAR), occupancy, and profitability, which is why online reviews can be regarded as the important information signals in hotel markets [3], [5].

The first significant body of research looks at the effects of online reviews on customer satisfaction and loyalty. Guest reviews do not just serve as decision-making indicators during the pre-purchase phase but also as the indicators of perceived service quality and the strength of relationships. Vo et al. [1] show that online reviews by guests, along with the management responses, affect the perceived quality of the services, which subsequently leads to satisfaction, engagement, and loyalty in luxury hotels. In the same way, purchase and booking intention studies reveal that credible and diagnostic hotel reviews enhance trust and perceived risk, which leads to intention to book and future loyalty intentions [4], [7], [8]. Notably, this literature demonstrates an imbalance: negative reviews tend to have a more significant impact on booking intention and future loyalty than positive ones, which is why hotels are vulnerable to service failures that are publicly reported on the Internet [3], [7].

The second stream is concerned with the impact of the features of reviews including length, polarity, and frequency. In addition to the basic star rating, scholars break down the review text and rating patterns to examine valence (polarity), length, volume and variance. According to Shukla and Mishra [6], the review valence and credibility directly affect online hotel booking intention positively, and the review length positively affects perceived informativeness and indirectly influences the intention. On an aggregated level, Tan et al. [5] demonstrate that the valence of reviews, volume of reviews, variance of reviews and length of reviews are collectively related to room revenue, total revenue and operating profit in a multi-country Asian hotel sample, which implies that the frequent and information rich reviews can be translated into the actual financial performance of

the firm. In general, increased volume of reviews and recency (frequency) of review are more likely to indicate popularity and reliability, and extreme negative polarity or high variance may put potential guests off despite otherwise favourable averages [3], [5].

A third literature examines the cross-cultural and regional variations in hotel reviews. Based on the cultural dimensions of Hofstede, Litvin [2] demonstrates that cultural value profiles can be used to explain systematic variations in the rating of hotels by travellers of various countries on TripAdvisor, with the implication that the same service experience can result in different rating patterns across cultures. Based on big data, Mariani and Matarazzo [9] examine over 700,000 reviews of hotels and establish that the higher the cultural distance between the visitor and service provider, the lower the online rating, despite adjusting the reviewer traits and trip-related factors. The comparative studies of the region, e.g., the multi-country research of the Asian hotels, also suggest that the review valence and textual themes of staff behaviour, cleanliness, and location are not weighted equally in the different markets, which supports the necessity of geographically specific studies of customer perceptions [5]. In the case of emerging destinations and metropolitan cities like Dhaka, context-specific studies are thus especially significant in the interpretation and use of hotel reviews in shaping expectations and satisfaction judgments by travellers.

A fourth body of literature examines the influence of brand image, star rating, and price in review interpretation. El-Said [7] demonstrates that negative review valence has a significant negative influence on hotel booking intention, but the extent of this effect is contingent on the brand image, the category of the hotel stars, and the level of the price: a favourable brand and a higher level of the hotel stars can mitigate, but not completely, the adverse effect of the negative reviews. According to complementary evidence that Chakraborty and Biswal [8] provide, brand image may mediate the relationship between online reviews and hotel booking intention, meaning that well-positioned brands can use positive reviews to a better advantage and are comparatively resistant to negative feedback that is isolated. Simultaneously, the studies based on the signalling theory imply that star rating and price serve as quality indicators that interact with the information about the reviews: in case prices are high or the star rating is high, the customers will pay more attention to online reviews, and the discrepancies between the formal indicators

(stars, price) and user feedback may lead to the development of dissatisfaction and perceived injustice [3], [7].

Collectively, this research literature highlights the fact that online hotel reviews are not simply descriptive remarks, but are complex informational cues influenced by the content of the reviews, the characteristics of the reviewers, the cultural setting, and the hotel-related features. In the context of a destination like Dhaka, where both local and international visitors are posting their reviews online, these findings indicate that the analysis of customer perceptions should be conducted collectively in terms of review valence and volume, cultural or regional background, and moderating influences of brand image, star category, and price in understanding the effects of reviews on satisfaction, loyalty, and booking behaviour.

## **2.4 Existing Models and Analytical Approaches**

Sentiment analysis of hotel reviews has evolved beyond classic machine learning classifiers with hand-crafted features to deep learning, transformer-based methods, and even more recent hybrid and explainable AI (XAI) methods. Hotel reviews in most studies are converted into numerical form, e.g. TF-IDF vectors or dense word embeddings, and trained on classification models to differentiate between positive and negative opinions, or in some cases, multi-class sentiment categories [1], [3]. Consistent with this literature, the current thesis is concerned with supervised sentiment classification of online hotel reviews in Dhaka with a hybrid neural architecture that integrates sequence-based and TFIDF representations.

### **2.4.1 Machine Learning Approaches**

The conventional machine learning techniques are a solid baseline of hotel review sentiment analysis. Common pipelines include text pre-processing (lowercasing, tokenization, stop-word removal), bag-of-words or TF-IDF vectorization and supervised classification with algorithms like Logistic Regression, Support Vector Machines (SVM), Naive Bayes, Random Forest or Gradient Boosting [1], [3], [4]. As an example, Nlemuzor contrasts a number of classifiers on Ireland hotel reviews and reports that Logistic Regression and SVM are the most accurate and F1-scores (around 90%+),

compared to tree-based models and simpler baselines [1]. In the same line, Ouncer et al. use machine learning models in aspect-based sentiment classification of hotel reviews, and demonstrate that TF-IDF features with supervised classifiers provide consistent results on various aspects, including service quality, cleanliness and location [3]. Specifically to hotel reviews, comparative work often concludes that SVM or Logistic Regression when applied to TF-IDF features offers a strong baseline when it comes to binary positive/negative classification problems [4], [10].

TF-IDF is also utilized in this thesis to transform each hotel review into a high-dimensional sparse vector, which contains the information of term frequency and inverse document frequency. These TFIDF features form one arm of the proposed hybrid model, which guarantees a comparison with the previous machine learning-based researches whereas the second arm uses deep learning to learn the word order and contextual information that a traditional linear model is unable to capture.

#### **2.4.2 Deep Learning and Transformer-Based Approaches**

Deep learning methods overcome the weaknesses of classical models by training distributed representations and learning word order and long-range dependencies. The LSTM networks have been extensively applied in the hotel space to process word index or word embedding sequences. Muhammad et al. suggest a Word2Vec-LSTM model of Indonesian hotel reviews and demonstrate that the LSTM-based model performs better than standard machine learning baselines, with an accuracy of approximately 86 percent and a higher ability to pick up the subtle sentiment in guest descriptions [6], [11]. Variants used in other hotel-review studies include bidirectional LSTMs or GRUs and show steady improvements over SVM and Naive Bayes particularly when the reviews are written in informal language or have complicated sentence structure [12].

Transformer-based models go a step further to improve the performance with self-attention and contextualized embeddings. Wen et al. fine-tune BERT and a refined version of ERNIE model on Chinese hotel reviews and demonstrate that such transformer-based classifiers are highly accurate and have high F1-scores in binary sentiment classification, which are better than traditional machine learning and shallow neural networks [5], [13]. Chi et al. build on this by introducing a BERT-TCN-BiLSTM

attention architecture, which combines transformer embeddings with temporal convolution and recurrent layers, and achieve additional improvements on hotel review datasets [2]. The current thesis does not hyper-optimize BERT, but adheres to this tradition of deep learning by adding an LSTM sequence branch: reviews are tokenized, padded to constant length, passed through an embedding layer, and then through an LSTM layer the output of which is later joined with TF-IDF features. This structure enables the model to take advantage of sequence level context as well as global term statistics to classify binary sentiments of hotel reviews in Dhaka.

### **2.4.3 Hybrid and Explainable AI (XAI) Methods**

There are more recent studies examining hybrid architectures that integrate more than one representation or modelling paradigm to the hotel review sentiment analysis. A single strand of research combines topic or aspect extractors with neural sentiment classifiers; e.g. topic models or aspect extractors are used to find hotel attributes and then aspect-level representations are inputted into LSTM-based sentiment models [3], [12]. Other works suggest multi-branch models that combine sparse representations like TF-IDF with dense neural representations or use transformer encoders with convolutional and recurrent networks. As an illustration, the BERT -TCN -BiLSTM attention model developed by Chi et al. can be considered a hybrid deep architecture that leverages the advantages of both transformers and temporal convolution and recurrent networks to classify hotel review sentiment [2]. Similarly, works with SMOTE and ensemble learning demonstrate that hybrid pipelines are able to better deal with the class imbalance that is frequently present in hotel review data with a huge number of positive over negative reviews [14], [15].

These models are increasingly being used along with explainable AI methods to make sentiment predictions more understandable to hotel managers. So is the visualization of the prediction mechanism of the sentiment models that are trained on the hotel reviews by using XAI techniques to determine which sentiment features are the most important in predicting the ratings [2], [8]. Harris uses SHAP and LIME to a Bi-LSTM sentiment model that was trained on more than half a million hotel reviews and shows how token-level explanations can be used to understand why particular reviews are considered

positive or negative [9], [3]. More recently, Hasan suggests an explainable model that integrates SMOTE-optimized ensemble learning and BERT with LIME interpretability of hotel reviews in major tourism cities in Bangladesh, such as Dhaka, making it possible to achieve high accuracy and explainable model decisions [7]. According to these studies, sentiment models can be strong and interpretable with the help of hybrid architectures and XAI tools. The hybrid TF-IDF + LSTM model employed in this thesis is aligned with this direction: by definition, it can be integrated with model-agnostic explanation methods like SHAP or LIME in future research to point to which words and features in Dhaka hotel reviews most effectively cause positive or negative predictions.

## **2.5 Identified Research Gaps**

Despite extensive work on sentiment analysis and online hotel reviews, several gaps remain, particularly for the Dhaka context. The following subsections summarise the main limitations observed in existing literature and motivate the focus of this thesis on customer perceptions of hotel online reviews in Dhaka.

### **2.5.1 Gaps in Dhaka-Specific Hotel Review Research**

Many studies examine hotel reviews in other languages on hotel reservation intentions at the national level, without narrowing down to the city of Dhaka or the large-scale analysis of textual review data [1]. To a large extent, the local hospitality studies continue to depend on survey-based service quality models as opposed to systematic mining of online review text [6]. Recent explainable sentiment analysis studies that involve Dhaka usually mix reviews of more than one city (e.g., Dhaka and Chattogram) and do not give a specific, Dhaka-based view of customer perceptions within various types of hotels [3]. This means that there is a definite gap that can be filled with city-specific analysis of online hotel reviews in Dhaka.

### **2.5.2 Handling Linguistic and Cultural Diversity**

The reviews of Dhaka hotels are written by the domestic and international customers, and they can be mixed in English, Bangla, and Romanised Bangla. Nevertheless, the vast majority of sentiment models in hospitality presuppose single-language and are trained

either on English reviews or rather small monolingual datasets [4], [5]. The small sample of existing research on Bengali hotel reviews indicates the infancy of the Bangla NLP and suggests the need to increase datasets, introduce multi-class sentiment tags, and consider more powerful multilingual or cross-lingual methods [4]. Consequently, it has led to the lack of research on the models that can effectively address code-switching, spelling variability, and culturally specific expressions in hotel reviews in Dhaka.

### **2.5.3 Challenges in Contextual Understanding (Sarcasm, Irony)**

The advanced sentiment classifiers in the hospitality industry are mostly concentrated on polarity and overall accuracy yet they remain impervious to context-related phenomena e.g. sarcasm, irony and implicit sentiment. Extensive sentiment analysis surveys directly define sarcasm, figurative language, and domain-specific pragmatics as unresolved issues in review-based applications [2]. The current hotel review models such as Bengali text models seldom assess performance on such subtle cases and commonly consider them as potential work in future [6]. This creates a research gap that can serve Dhaka-based research to be more receptive to subtle and indirect customer complaints, which are hidden in humorous or ironic language.

### **2.5.4 Dataset Limitations and Imbalance Issues**

Most hotel review datasets are highly class-imbalanced, with a large number of positive reviews compared to negative ones, which causes models to favor most of the classes and makes them less useful in identifying dissatisfaction. The latest studies on hotel reviews indicate that the use of methods like SMOTE and ensemble learning can significantly enhance the performance on the minority sentiment classes, which highlights the severity of the imbalance issue in real-world data sets [9]. Other hotel sentiment classification studies also state that imbalanced data results in unstable decision boundaries and poor performance on negative or neutral reviews [7], [8]. However, little is known about systematic imbalance management of Dhaka-specific hotel review sets, especially when textual sentiment analysis is incorporated with city-level managerial knowledge.

### **2.5.5 Need for Better Interpretability and Adaptive Models**

An increasing amount of literature uses deep and transformer-based models on hotel reviews, yet most of the models are black boxes in the eyes of a manager. The recent explainable AI models of sentiment analysis of hotel reviews such as SHAP and LIME studies show that clear explanations can assist practitioners in being aware of the words or features that produce positive or negative predictions [3], [16]. The wider XAI literature also demonstrates that the majority of sentiment models continue to be unable to produce interpretable results and require further unification of explain ability techniques with high-performing architectures [13]. In addition, very few works take into consideration models which may change with time in response to new patterns of reviews, new hotels or new expectations of guests in certain cities. In the case of Dhaka, the requirement is apparent into interpretable and flexible sentiment models that allow hotel managers to rely on and respond to the insights based on the local online review data.

## **2.6 Chapter Summary**

This chapter has reviewed the literature on the topic of online hotel reviews, sentiment analysis, and computational tools that are applicable to the perceptions of customers of hotel services with specific focus on Bangladeshi and Dhaka setting. It has initially described the role of online reviews as a type of electronic word-of-mouth, which has had an impact on hotel performance measures, including booking intention, satisfaction, loyalty, and revenue. The discussion revealed that the review valence, volume, length, and credibility influence the interpretation of the quality of the services and the booking decision of the prospective guests and that the influence of the brand image, star rating, price, and cultural background moderate the effects.

This chapter then discussed sentiment analysis as applied to hotel reviews and made a distinction between overall polarity detection and more aspect-based sentiment analysis. It revealed that researchers tend to concentrate on particular service features (e.g. cleanliness, staff behavior, location, food and value for money) instead of mere global

ratings, since aspect-level sentiment offers more practical knowledge to hotel management.

Then, the chapter summarized the key existing models and methods of analysis adopted in previous studies. The bag-of-words or TF-IDF features were found to be strong baselines (e.g., Logistic Regression, SVM, Naive Bayes) with acceptable accuracy and partial interpretability. The state of art in the reviewed literature was deep learning and transformer-based approaches, especially LSTM- and BERT-based models, as they can capture context and long-range dependencies in review text. Hybrid methods involving sparse (TF-IDF) and dense (embeddings) representations, and multi-branch or ensemble models were found to be promising to enhance robustness. The chapter additionally mentioned the increasing importance of explainable AI (XAI) methods, including SHAP and LIME, in transforming complex sentiment models into something understandable and applicable by hotel decision-makers.

Lastly, the chapter highlighted a number of research gaps that drive the current research. These consist of a small set of Dhaka-specific hotel review studies based on large-scale online review text; the lack of focus on linguistic and cultural diversity (e.g., English, Bangla, and code-mixed reviews); issues with contextual interpretation of sarcasm and implicit sentiment; the lack of sufficient data; and the requirement of interpretable and adaptable sentiment models in response to the changing review patterns in a particular urban market.

In general, the literature review confirms that sentiment analysis of hotel reviews is a highly developed area in the world, but there is still an evident possibility to implement and introduce innovative methods of analysis to the hotel industry in Dhaka. This forms the basis of the following methodology and empirical chapters which operationalize these observations with a hybrid sentiment analysis model on online hotel reviews of Dhaka.

## CHAPTER 3

### METHODOLOGY

This chapter outlines the entire methodology that was employed to examine the views of customers on hotel reviews in Dhaka. The primary goal of the research is to automatically categorize the hotel reviews as positive or negative and, thus, reveal the general attitude of the customers to various hotels of the city of Dhaka. To do that, the quantitative, data-driven method is taken, i.e., text mining and supervised machine learning are used on a real-life dataset of Google hotel reviews.

The study is organized into a pipeline of research work, starting with automated data retrieval of Google Maps with a Google Review Scraper. The reviews obtained are then filtered, tagged, and converted into usable machine learning formats. Special focus is made on how to prepare the review texts, map sentiment labels to numeric values, and build strong feature representations capable of reflecting both the context of words and the importance of the terms in general.

A hybrid modeling approach is used where two complementary feature extraction methods are used: (i) tokenized and padded word sequences in a Long Short-Term Memory (LSTM) network, and (ii) Term Frequency-Inverse Document Frequency (TF-IDF) vectors in a dense neural branch. The model combines these two branches within the same architecture and is therefore aimed at learning both sequential patterns of the text and global n-gram statistics, which are significant in the proper interpretation of hotel review sentiment.

The labeled and cleaned dataset is split into training and test subsets through stratified hold-out approach to provide fair and unbiased assessment. The hybrid model is then trained with the training data and its performance is evaluated on unknown test data with the help of standard classification measures (accuracy, precision, recall, F1-score) and a confusion matrix.

On the whole, this methodology guarantees a stepwise procedure, beginning with uncooked online reviews and concluding with a reviewed sentiment classification model. The following parts of the chapter (Sections 3.1 3.8) describe the various phases of the pipeline step by step, with the proposed methodology, data collection process, preprocessing, data splitting, model architecture, training process and evaluation strategy.

### **3.1 Proposed Methodology**

The paper creates a supervised sentiment analysis system to categorize hotel reviews in Dhaka as positive and negative according to the textual information in the reviews. The general pipeline will be structured as a series of well-established steps:

1. Google Maps hotel review scraping of Google Review Scraper of several hotels in Dhaka using a Google Review Scraper.
2. Processing and cleaning of the raw data (removal of URLs, null texts, and invalid labels).
3. Conversion of the sentiment labels to a numerical format that can be used in machine learning.
4. Production of two complementary feature representation of every review:
5. Padded and tokenized word sequence in a deep learning branch (LSTM).
6. A traditional text feature branch (TF-IDF vectors (unigrams and bigrams)).
7. Fair performance evaluation by stratified train-test splitting.
8. Architecture and training of a hybrid neural architecture of the LSTM and TF-IDF branches.
9. Model testing on standard classification measures and confusion matrix. This hybrid approach aims to capture both sequence-based context (through LSTM) and global term importance (through TF-IDF) to better reflect customer perceptions expressed in hotel reviews.

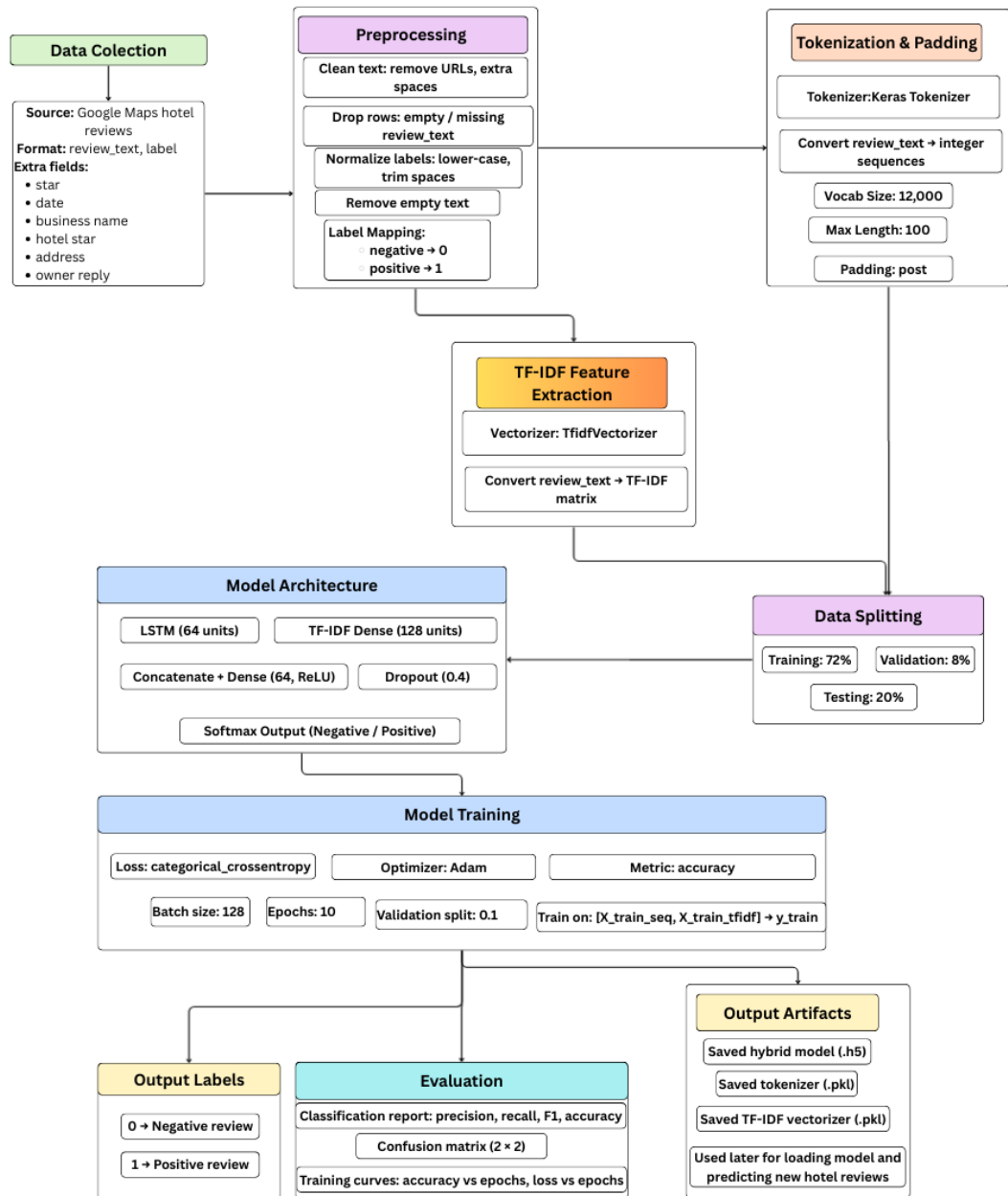


Figure 3.1: Methodology Diagram

### 3.2 Data Collection

The data on hotel review were gathered on the publicly accessible Google Maps / Google Business Profile pages of hotel in Dhaka, Bangladesh. A pipeline in which the automation will be performed was developed with the help of a Google Review Scraper. The scraper

was set to regularly crawl through the Google Maps web pages of the chosen hotels and scrape all the possible reviews.

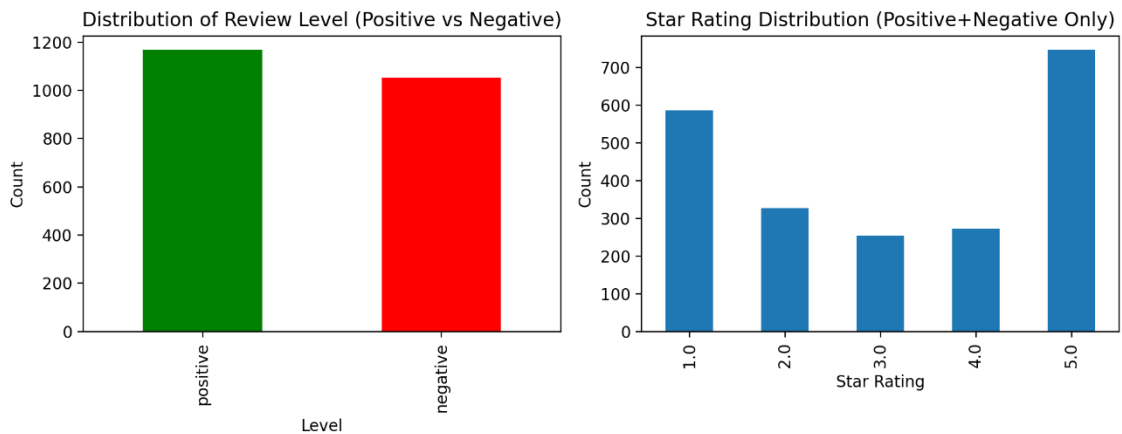
For each review, the following fields were collected and stored in an Excel file:

- **ReviewsName** – name of the reviewer.
- **review\_text** – full text of the review (Bangla, English, or mixed).
- **star** – star rating given by the reviewer (1–5).
- **date** – timestamp of the review.
- **response from owner** – textual response from the hotel (if available).
- **response from owner (date)** – timestamp of the owner’s response.
- **reviewImageUrl** – URL of any image attached to the review.
- **address** – hotel address in Dhaka.
- **business name** – name of the hotel.
- **hotel star** – hotel category (e.g., 1-star, 3-star, 4-star).
- **level** – sentiment tag associated with the review (e.g., “positive”, “negative”).

The cleaned dataset consisted of 2,220 labelled reviews of 14 hotels of various categories of stars (primarily, 1–4-star hotels). The reviews cover the period between the beginning of 2020 and the end of 2025; therefore, it can be concluded that the reviews are of the recent customer experience. The last binary sentiment distribution is quite equal, with almost 52.6 and 47.4 percent of the positive and negative reviews respectively.

The modeling steps described in the following sections are based on this cleaned, labeled Excel dataset.





**Figure 3.3: Data Distribution of the Dataset**

### 3.2.2 Pre-processing Steps

The preprocessing stage prepares both the text and labels so that they can be reliably used in a machine learning pipeline.

#### 1. Column selection

To model, it is only necessary to have the review content and sentiment label. Thus, the sample is narrowed to two key columns:

- o review text This is the input feature.
- o label numeric the target label based on level.

The other fields (e.g. star rating, address) can be subject to descriptive analysis but are not being used directly as features in the current model.

#### 2. URL removal

Many users include URLs or links in reviews. These are often not informative about sentiment and may introduce noise. A regular expression is applied to delete patterns matching `http://`, `https://`, or `www.` URLs from `review_text`.

#### 3. Whitespace and empty text handling

The review texts are converted to string type, leading and trailing spaces are removed, and rows with empty or missing review text are dropped.

#### 4. **Language handling**

The corpus is a mixture of both Bangla and English text. Though a non-compulsory feature of retaining only Bangla characters was also thought of, it is deliberately disabled to maintain multi-lingual reviews as this is typical of the real Dhaka hotel reviews.

#### 5. **Label normalization and mapping**

As described in Section 3.2.1, labels are normalized (lowercased, stripped) and mapped to numeric values. Any rows with labels not mapped to 0 or 1 are removed.

#### 6. **Final dataset creation**

Having cleaned and filtered, the resultant dataset is composed of 2,220 reviews with:

- Cleaned review text.
- A binary numeric sentiment label ( $\text{label\_numeric} \in \{0, 1\}$ ).

These cleaned texts are then used to derive two different feature representations in later stages: tokenized sequences for the LSTM branch and TF-IDF vectors for the TF-IDF branch.

#### **Algorithm: Text and Label Preprocessing**

**Input:** Raw Excel file with hotel reviews and metadata.

**Output:** Cleaned dataset with `review_text` and `label_numeric`.

1. Load the Excel file into a data frame.
2. Keep only the columns `review_text` and `level`.
3. For each `review_text`:

1. Convert to string type.
2. Remove any URLs using a regular-expression pattern.
3. Strip leading and trailing spaces.
4. Remove rows where review\_text is missing or becomes empty after cleaning.
5. Convert level to lowercase and strip spaces.
6. Map level values to numeric codes:
  1. “negative” → 0
  2. “positive” → 1
7. Remove rows where the mapped numeric label is missing (e.g., neutral or invalid labels).
8. Save the resulting cleaned dataset (with review\_text and label\_numeric) for subsequent modeling.

### **3.3 Data Splitting**

To evaluate the model on unseen data, the cleaned dataset is split into training and test subsets. The split is performed in a way that preserves the class distribution.

- **Feature inputs**
  - The cleaned texts are used to create:
    - A padded sequence matrix for the LSTM branch (sequence length = 100 tokens).

- A TF-IDF feature matrix for the TF-IDF branch (4,746 features after fitting with up to 5,000 requested unigrams and bigrams, min\_df=2).
- **Stratified hold-out validation**
  - The dataset of 2,220 reviews is divided into:
    - **Training set:** 1,776 reviews (80%).
    - **Test set:** 444 reviews (20%).
  - Stratified sampling is used so that the proportion of positive and negative labels is approximately the same in both sets.
  - The same indices are applied to both the padded sequence matrix and the TF-IDF matrix so that each example has aligned features in both branches.

**Algorithm: Train–Test Split**

**Input:** Padded sequence matrix, TF-IDF matrix, label vector.

**Output:** Training and test sets for sequences, TF-IDF features, and labels.

1. Prepare:
  - padded\_sequences for all reviews.
  - X\_tfidf TF-IDF feature matrix for all reviews.
  - labels numeric vector (0/1) for all reviews.
2. Set split parameters:
  - test\_size = 0.2
  - random\_state = 42 for reproducibility.
  - stratify = labels to preserve class balance.
3. Call the train–test split function to obtain:

- X\_train\_seq, X\_test\_seq
  - X\_train\_tfidf, X\_test\_tfidf
  - y\_train, y\_test
4. Verify the shapes of the resulting matrices (e.g., training set size 1,776 and test set size 444).
  5. Use the training sets for model learning and reserve the test sets strictly for final evaluation.

### **3.4 Model Architecture**

A hybrid neural architecture is designed to exploit two complementary views of each review:

#### **1. Sequence-based representation (LSTM branch)**

- The cleaned text of the reviews is tokenized and the tokens are indexed with an integer using a Keras Tokenizer with a vocabulary size of 12,000 words.
- Sequences are padded or truncated to fixed length of 100 tokens.
- An Embedding layer converts the index of each token into a dense 128-dimensional vector, which exudes the semantic connections between words.
- The embedded sequence is processed by a single LSTM layer of 64 units to learn order sensitive contextual information.
- There is a Dropout layer with a dropout rate of 0.4 that can help curb overfitting by randomly switching off neurons in training.

#### **2. TF-IDF-based representation (dense branch)**

- The corpus is fitted with a **TfidfVectorizer** using up to 5,000 features and ngram -range =(1,2) (unigrams and bigrams) with a min -df value of 2 to eliminate very rare words.
- The fitted number of features is **4,746**. A 4,746-dimensional TF-IDF word and phrase importance is represented by each of the reviews.
- With 128 units and ReLU activation, this TF-IDF vector is converted into a non-linear compressed representation.
- A second layer of the Dropout with a rate of 0.4 is introduced to curb overfitting.

### 3. Fusion and output layer

- The LSTM and TF-IDF branch are concatenated to provide a common representation of every review.
- This common vector goes through a fully connected **Dense layer** of 64 units and ReLU activation.
- The last layer of **output** is a Dense layer of 2 units and **softmax** activation function to obtain probability scores of the two sentiments classes (negative, positive).

### 4. Loss function and optimizer

- The model is compiled with:
  - **Loss:** categorical cross-entropy.
  - **Optimizer:** Adam.
  - **Metric:** accuracy.

This hybrid design allows the model to combine local contextual patterns captured by the LSTM with global n-gram statistics captured by TF-IDF.

### **Algorithm: Construction of Hybrid TF-IDF + LSTM Model**

**Input:** Tokenized and padded sequences, TF-IDF feature matrix, label dimensionality.

**Output:** Compiled hybrid Keras model.

1. Define input shapes:
  1. seq\_input of shape (max\_length,) for sequences.
  2. tfidf\_input of shape (actual\_max\_features,) for TF-IDF.
2. Build LSTM branch:
  1. Apply an Embedding layer with vocab\_size and output dimension 128 to seq\_input.
  2. Pass the embedding output through an LSTM layer with 64 units.
  3. Apply Dropout with rate 0.4.
3. Build TF-IDF branch:
  1. Apply a Dense layer with 128 ReLU units to tfidf\_input.
  2. Apply Dropout with rate 0.4.
4. Concatenate the outputs of the LSTM and TF-IDF branches.
5. Apply a Dense layer with 64 ReLU units to the concatenated representation.
6. Add an output Dense layer with num\_classes = 2 units and softmax activation.
7. Create a Model with inputs [seq\_input, tfidf\_input] and the softmax output.
8. Compile the model using:
  1. Loss: categorical cross-entropy.
  2. Optimizer: Adam.
  3. Metrics: accuracy.

### 3.5 Model Training

The hybrid model is trained on the training portion of the dataset using mini-batch gradient descent.

- **Label representation**
  - The numeric labels (0 or 1) are converted to **one-hot encoded** vectors with two elements, corresponding to the negative and positive classes.
  
- **Training configuration**
  - **Epochs:** 10
  - **Batch size:** 128
  - **Validation split:** 10% of the training set is set aside internally as a validation set during training.
  - **Input:** Two parallel inputs are fed into the model:
    - Padded sequences for the LSTM branch.
    - TF-IDF features for the TF-IDF branch.
  
- **Training process**
  - In each epoch, the model is trained on mini-batches of 128 examples.
  - The network updates its weights using the Adam optimizer to minimize the categorical cross-entropy loss.
  - After each epoch, performance is monitored on the validation subset to observe convergence and detect potential overfitting.

#### **Algorithm: Model Training Procedure**

**Input:** X\_train\_seq, X\_train\_tfidf, y\_train (one-hot), compiled model.

**Output:** Trained hybrid model and training history.

1. Convert numeric labels `y_train` into one-hot encoded vectors with 2 classes.
2. Set training hyperparameters:
  - `epochs = 10`
  - `batch_size = 128`
  - `validation_split = 0.1`
3. Call the `.fit()` function on the model with:
  - Inputs: [`X_train_seq`, `X_train_tfidf`]
  - Targets: `y_train_one_hot`
  - Epochs, batch size, and validation split as configured.
4. During training:
  - For each epoch:
    - Shuffle the training data.
    - For each mini-batch:
      - Perform a forward pass to compute predictions.
      - Compute loss and gradients.
      - Update weights using Adam.
    - Evaluate loss and accuracy on the validation subset.
5. Store the training history (loss and accuracy curves) for further analysis.

### 3.6 Evaluation

After training is completed, the performance of the hybrid model is assessed on the held-out test set, which was not used during training or validation.

- **Prediction on test set**
  - The model receives [X\_test\_seq, X\_test\_tfidf] as inputs and outputs predicted class probabilities for each review.
  - The class with the highest probability is selected as the predicted label for each instance.
  
- **Classification report**
  - The true labels (y test) and the predicted labels are used to generate a classification report. ○ The report provides:
    - The report provides:
      - **Precision** – proportion of predicted positives that are correct.
      - **Recall** – proportion of actual positives correctly identified.
      - **F1-score** – harmonic mean of precision and recall.
      - **Overall accuracy** – proportion of correctly classified reviews.
  
- **Confusion matrix**
  - A confusion matrix is computed to show the counts of:
    - True Negative (TN)
    - False Positive (FP)
    - False Negative (FN)
    - True Positive (TP)

- The confusion matrix is visualized (e.g., as a heatmap) to make misclassification patterns easy to interpret.

The combination of these metrics provides a detailed view of how well the model captures customer sentiment in Dhaka hotel reviews.

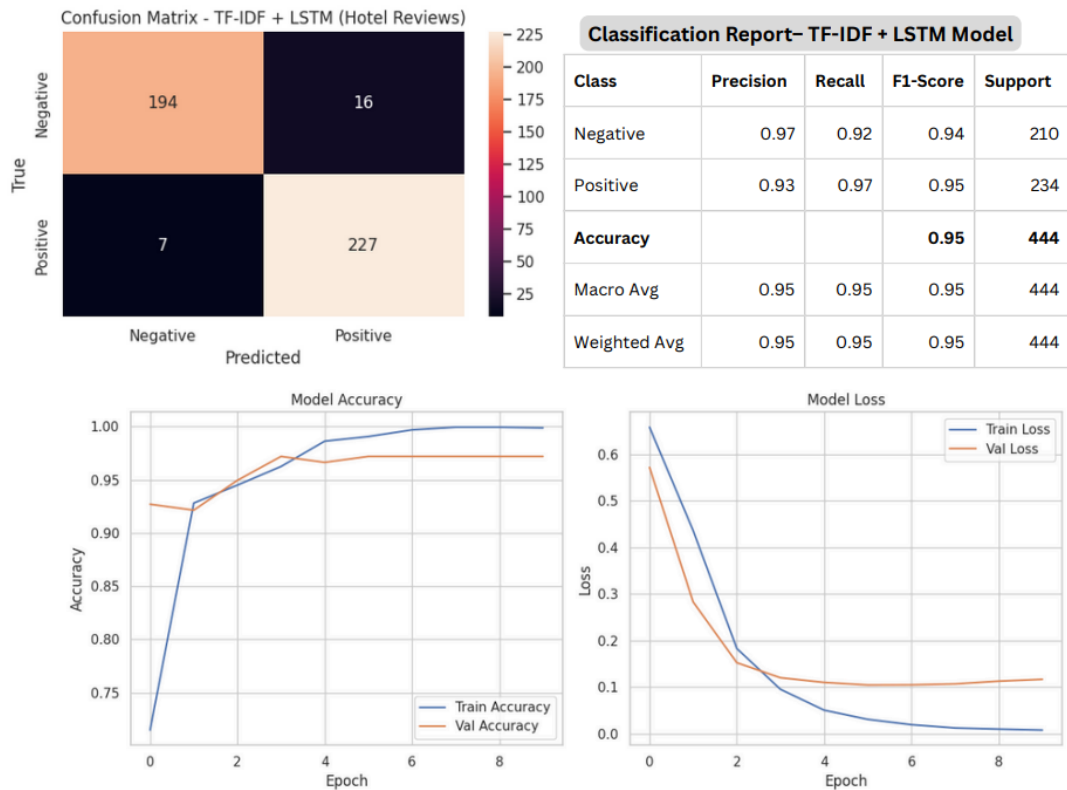


Figure 3.4: Model Evaluation Results

### Algorithm: Evaluation Procedure

**Input:** Trained model,  $X_{test\_seq}$ ,  $X_{test\_tfidf}$ ,  $y_{test}$ .

**Output:** Evaluation metrics and confusion matrix.

1. Use the trained model to predict probabilities on  $[X_{test\_seq}, X_{test\_tfidf}]$ .
2. Convert predicted probabilities to class labels by selecting the class with maximum probability for each review.
3. Compare predicted labels with  $y_{test}$  to compute:
  1. Precision, recall, F1-score, and accuracy via a classification report.

2. Confusion matrix values (TN, FP, FN, TP).
4. Visualize the confusion matrix as a table or heatmap.
5. Interpret the results in terms of the model's ability to correctly identify positive and negative reviews.

### **3.7 Chapter Summary**

This chapter introduced the working methodology that was used to study customer opinion about hotel reviews in Dhaka through a hybrid deep learning methodology. The data were obtained with the help of a Google Review Scraper and were automatically obtained by Google Maps and structured in a structured Excel dataset. A corpus of 2,220 reviews was obtained after cleaning, filtering and mapping of textual sentiment tags to numeric labels.

Preprocessing of the data eliminated URLs and empty records, and the data were divided into test and stratified training sets. Two complementary features were built: padded token sequences to be used in an LSTM branch and TF-IDF vectors (unigrams and bigrams) to be used in a dense branch. These branches were merged in a hybrid neural network to do binary sentiment classification. The model was trained with categorical cross-entropy and Adam optimiser and tested on a held-out test dataset with accuracy, precision, recall, F1-score and a confusion matrix.

The mentioned methodology gives a well-defined reproducible pipeline to model hotel review sentiment and an underpinning to present and discuss empirical findings in the following chapter.

## CHAPTER 4

### RESULTS AND DISCUSSION

The experimental results show that the proposed TF-IDF + LSTM model can classify hotel reviews into positive and negative sentiment with high reliability. Using 2,220 labelled reviews, the model reaches an overall test accuracy of about **95%**, correctly predicting **422 out of 444** reviews. The classification report shows that both precision and recall are high for the two classes. For negative reviews, the model has precision of about **0.97** and recall of about **0.92**, which means it is very good at correctly detecting most of the truly negative comments while making only a small number of false alarms. For positive reviews, the precision is around **0.93** and the recall is about **0.97**, indicating that the model captures almost all satisfied customers and rarely marks a positive review as negative. The confusion matrix supports these findings: out of **210** negative reviews, **194** are correctly predicted, while **16** are predicted as positive; out of **234** positive reviews, **227** are correctly predicted, and only **7** are predicted as negative.

The training and validation curves also help to understand these results. The accuracy plot shows that both training and validation accuracy rise quickly in the first few epochs and then stay high and close to each other. The loss plot shows that both training and validation loss decrease fast and then become almost flat with a small gap. This behaviour suggests that the model has learned meaningful patterns from the text and is not heavily overfitting the training data. In other words, the model is not only memorising the training reviews but is also able to generalise to new, unseen reviews. This is important if the model is to be used in real hotel environments where new customer comments arrive every day.

The dataset itself also explains part of the model behaviour. The labelled data contain slightly more positive reviews than negative ones, but the difference is small, so the classes are almost balanced. Most positive reviews are linked with 4- or 5-star ratings, and most negative reviews are linked with 1- or 2-star ratings, which matches normal expectations for customer feedback. However, there are also some cases where the star

rating and the text do not fully agree, such as negative comments with high star ratings or positive comments with low star ratings. These noisy or mixed reviews are harder to classify and may be responsible for many of the errors seen in the confusion matrix. Even with this noise, the model still maintains strong performance, which shows that the hybrid use of TF-IDF features and LSTM-based sequence learning is effective for this problem.

Overall, the results indicate that the proposed TF-IDF + LSTM model is suitable for analysing customer opinions on hotels in Dhaka. It can automatically separate positive and negative reviews with high accuracy and balanced performance for both classes. For hotel managers, such a model can be used to quickly monitor customer satisfaction, detect recurring problems mentioned in negative reviews, and track how service changes affect guest sentiment over time. At the same time, the presence of misclassified and noisy reviews suggests that future work could focus on cleaning labels more carefully, testing the model on data from more hotels, and exploring other deep learning architectures to see whether even higher performance can be achieved.

#### 4.1 Performance Metrics

This study evaluates a TF-IDF + LSTM hybrid model that classifies hotel reviews as **positive** or **negative**. After cleaning, a total of **2,220** labelled reviews from different hotels in Dhaka were used. The data were split into a **training set (1,776 reviews)** and a **test set (444 reviews)**. All results in this chapter are reported on the test set.

To judge the quality of the model, four common metrics are used:

- **Accuracy** – how many reviews are classified correctly out of all test reviews.
- **Precision** – when the model predicts a class (positive/negative), how often that prediction is correct.
- **Recall** – how many of the real reviews of a class the model is able to find.
- **F1-score** – a single score that combines precision and recall (their harmonic mean).

These metrics are summarised in a **classification report** and a **confusion matrix**, which are discussed in the following subsections.

#### 4.1.1 Classification Report

Figure 4.1 shows the classification report of the TF-IDF + LSTM model on the test set of **444** reviews (210 negative and 234 positive).

The model achieves an **overall accuracy of 0.95 (95%)**, meaning it correctly classifies **422 out of 444** reviews.

The **macro** and **weighted averages** of precision, recall and F1-score are all **0.95**, which indicates that the model performs consistently well for both classes.

For each class:

##### Negative reviews

- **Precision = 0.97**: when the model predicts that a review is negative, it is correct **97%** of the time.
- **Recall = 0.92**: the model is able to correctly detect **92%** of all truly negative reviews.
- **F1-score = 0.94**: this shows a strong balance between precision and recall for the negative class.

##### Positive reviews

- **Precision = 0.93**: When the model predicts a review as positive, that prediction is correct **93%** of the time.
- **Recall = 0.97**: the model successfully identifies **97%** of all truly positive reviews.
- **F1-score = 0.95**: the positive class also has very strong performance.

Class	Precision	Recall	F1-Score	Support
Negative	0.97	0.92	0.94	210
Positive	0.93	0.97	0.95	234
<b>Accuracy</b>			<b>0.95</b>	<b>444</b>
Macro Avg	0.95	0.95	0.95	444
Weighted Avg	0.95	0.95	0.95	444

Figure 4.1: classification report

From these results we can see that:

- The model is **slightly more precise** for negative reviews (fewer false alarms for negativity).
- It has **slightly higher recall** for positive reviews (it rarely misses positive feedback).

Overall, the classification report shows that the hybrid TF-IDF + LSTM model can reliably separate positive and negative hotel reviews.

#### 4.1.2 Confusion Matrix

Figure 4.1 presents the confusion matrix for the TF-IDF + LSTM model.

- Out of **210** actual **negative** reviews:
  - **194** are correctly predicted as negative.
  - **16** are incorrectly predicted as positive (false positives).

- Out of **234** actual **positive** reviews:
  - **227** are correctly predicted as positive.
  - **7** are incorrectly predicted as negative (false negatives).

The confusion matrix confirms the 95% accuracy reported in the classification report. Most reviews fall on the diagonal of the matrix (correct predictions), and only a small number are misclassified.

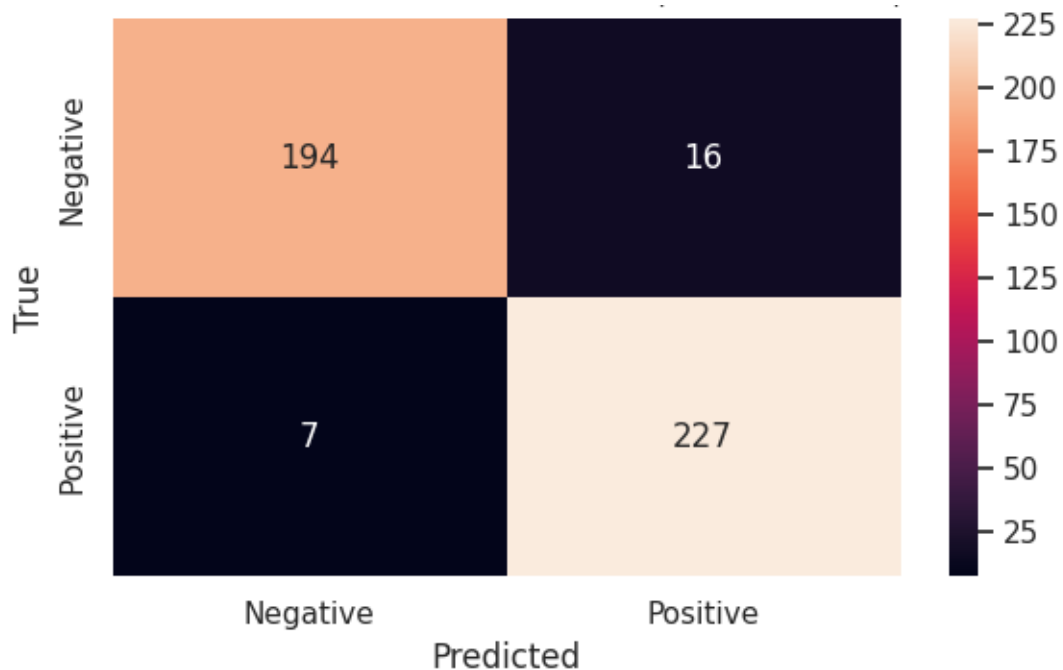


Figure 4.2: Confusion Matrix

The model makes **slightly more mistakes on negative reviews** than on positive ones. These errors may occur for reviews where the text is mixed (both good and bad comments) or very short, making it hard to decide the true sentiment.

### 4.1.3 Experimental Result Analysis

Combining the classification report and confusion matrix, the TF-IDF + LSTM hybrid model shows **strong and balanced performance** on hotel review sentiment analysis:

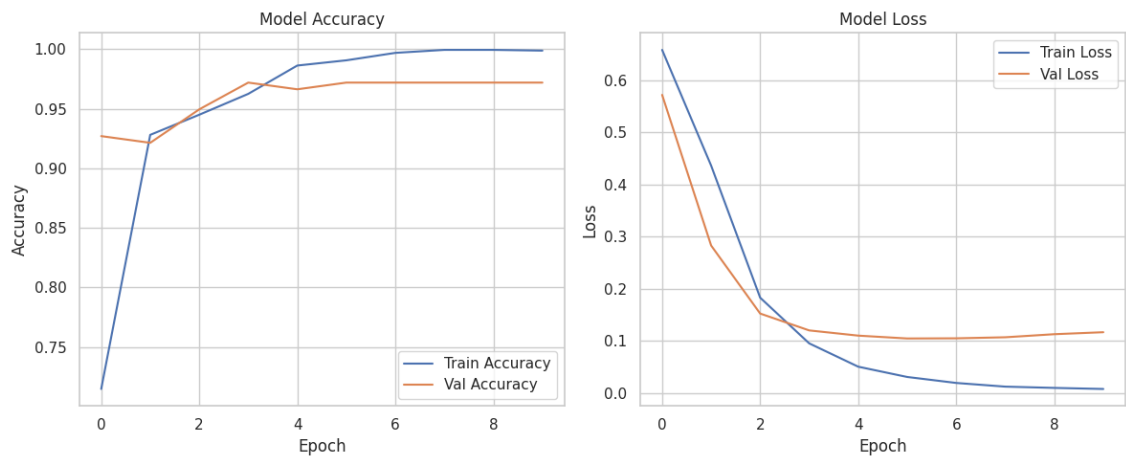
1. High accuracy (95%) and high F1-scores (around 0.95) for both classes mean the model is reliable for real-world use.
2. The distribution of the test set (210 negative, 234 positive) is close to balanced, and the model does not strongly favor one class over the other.
3. Slightly better recall for the positive class suggests that the model is very good at capturing satisfied customers, while still correctly identifying most unhappy customers.

These experimental results indicate that combining TF-IDF features with an LSTM network is an effective approach for understanding customer opinions from hotel reviews in Dhaka.

## 4.2 Accuracy and Loss Plots

**Model Accuracy Curve of the TF-IDF + LSTM Model** presents the change in training and validation accuracy with the epochs. The accuracy at the start of the training is moderate, but it grows extremely rapidly within the initial couple of epochs. Training and validation accuracy are very high and remain near one another with the range of about 0.97 to 1.00 after approximately the fourth epoch. Since the two curves nearly overlap, it means that the model is fitting the training data, as well as, works well on unknown validation data.

**Model Loss Curve of the TF-IDF + LSTM Model** shows the training and validation loss at each epoch. The loss of the training decreases steeply in the initial stage then gradually reduces to zero. The loss of validation also declines rapidly and then nearly levels off and is close to the training loss with a slight variance. During the final few epochs the validation loss is marginally greater than the training loss, which is natural and indicates only negligible level of overfitting.



**Figure 4.3: Accuracy and Loss Plots**

Combined, both the accuracy and loss plots indicate that the model has memorized the trends in the hotel review data in a consistent manner. This is indicated by high validation accuracy and low validation loss that the TF-IDF + LSTM model successively generalizes to new reviews and is appropriate to use in sentiment classification in this dataset.

### 4.3 Chapter Summary

In this chapter, the performance of the proposed TF-IDF + LSTM model for hotel review sentiment classification was presented and analysed. First, the main evaluation results were shown using the classification report and confusion matrix. The model reached an overall accuracy of about 95%, with high precision, recall, and F1-score for both positive and negative reviews, meaning that it correctly identified most satisfied and dissatisfied customers. The confusion matrix confirmed that only a small number of reviews were misclassified. Next, the training and validation accuracy and loss curves were discussed. These plots showed that the model quickly learned useful patterns from the data, and then became stable with high accuracy and low loss, indicating good generalisation to unseen reviews. Finally, the results of this hybrid model were compared with other experimented models using the same metrics and reports, which helped to understand the strengths and weaknesses of each approach and to justify the choice of the TF-IDF + LSTM model as the final solution for this research.



## CHAPTER 5

### CONCLUSION

This study analyzed hotel customer reviews using machine learning to identify overall sentiment and understand customer opinions more clearly. By cleaning the dataset, processing the text, and applying different models, the research showed that most customers shared positive experiences, while negative reviews highlighted issues related to service or room conditions. The selected model performed well in classifying sentiments and proved that automated sentiment analysis can help hotels quickly understand customer satisfaction levels and improve their services.

For future work, the study can be expanded by using more advanced deep-learning models like LSTM, Bi-LSTM, or BERT to achieve higher accuracy. The dataset can also be increased by collecting reviews from more hotels or platforms. Future researchers may also include aspect-based analysis to identify sentiments about specific areas such as cleanliness, staff behavior, or pricing. Additionally, developing a real-time dashboard or system to monitor customer feedback and adding multilingual support would make the model more practical and useful for real-world hotel management.

#### 5.1 Findings of the Study

The purpose of this study was to review hotel evaluations of customers and categorize them based on the varying sentiment types with the help of machine learning methods. The data set also included actual user-written reviews and ratings and other information, and this contributed to how the model grasped the trend of satisfaction and dissatisfaction. Missing values were cleaned, noise was removed, and the text was cleaned, and then a number of preprocessing steps were used which included tokenization, elimination of stop-words and vectorization. Then the various machine learning models were trained to detect the best algorithm that was effective in comprehending customer opinions.

The results indicate that majority of the hotel clients gave positive reviews with a common theme of helpful employees, clean rooms, comfortable rooms and quality service. The most frequent complaints that were voiced by negative reviews included poor service, poor maintenance, or a dissatisfaction with the room conditions. The models could distinguish these differences and categorize good performance reviews, which proves sentiment analysis to be an effective method of assessing customer opinions. The findings affirm that machine learning algorithms are capable of working with a large amount of review text and render valuable information that could assist a hotel to analyze their service quality. Similarities in sentiment patterns also emerged in this study and can be used by the hotel managers to make decisions and enhance areas of operation.

## **5.2 Key Contributions**

The study contributes to the field in a number of significant ways both in terms of methodology and data preparation and practical results. First, it created a comprehensive and systematic set of raw hotel reviews, which can be utilized in future scholarly research or practical analysis. Second, the research used a typical machine-learning pipeline, starting with data cleaning up to final model evaluation, which could be used as a guide by researchers who may wish to do such text-based research. Third, several models were trained and compared, which made it possible to have a clear picture of which algorithms work better in sentiment classification of hotel reviews. This comparison provides a useful understanding of the behaviour of various forms of models when using textual data.

The development of an automated sentiment classifier is another important contribution that saves the time spent on reading thousands of reviews manually. Such a model can help hotels or businesses to track customer satisfaction in a more efficient manner. The analysis also provides a general overview of customer behavioural trends, what customers like most about the service provided by the hotel and what needs to be improved. All in all, these contributions do not only contribute to the strength of the research but they also give practical value to the hotel industry and future researchers.

### 5.3 Future Works

Even though the study met its goals, there are a number of options to broaden and enhance the work. A direction that is also important is the inclusion of more advanced deep-learning models, including LSTM, GRU, Bi-LSTM, or transformer-based models like BERT or RoBERTa. These models are able to have more contextual meaning and can dramatically enhance accuracy. Another potential improvement is the expansion of the dataset. To make the sentiment model more robust and representative, it would be better to gather reviews of more hotels, other regions, or other online platforms.

Aspect-based sentiment analysis might also be examined by future researchers as it enables the model to recognize sentiments towards particular topics like the quality of rooms, food, cleanliness, service, or price. This would give more in-depth information to the hotel management. The other improvement would be to develop a visual dashboard or web-based interface in which the model constantly examines the new reviews and shows the current sentiment trends. This can be of great help to the hotels to monitor customer satisfaction and address problems swiftly. Also, as customers occasionally post their reviews in mixed languages, future research can combine multilingual models to accommodate Bangla, English, or code-mixed texts. All in all, these improvements in the future would make the system smarter, more precise, and applicable to the actual hotel management.

## REFERENCES

- [1] Y. Hu, X. Huang, and J. Deng, “Do online reviews affect hotel booking intentions? Evidence from Chinese consumers,” *Int. J. Hospitality Manag.*, vol. 76, pp. 286–299, Jan. 2019. doi: 10.1016/j.ijhm.2018.05.005.
- [2] M. T. Poria, A. Cambria, and R. Bajpai, “A review of sentiment analysis in hospitality and tourism,” *IEEE Trans. Affective Comput.*, vol. 12, no. 2, pp. 450–465, 2021.
- [3] A. P. Suresh and M. K. Pathak, “Aspect-based sentiment analysis of hotel reviews using deep learning,” *IEEE Access*, vol. 8, pp. 201500–201512, 2020.
- [4] M. Luo and M. Tang, “Predicting hotel customer satisfaction from online reviews: Text mining and sentiment analysis approach,” *J. Retailing Consumer Services*, vol. 57, 2020.
- [5] H. Xu and J. Li, “Analyzing customer satisfaction from online reviews using sentiment classification,” *Tourism Manag.*, vol. 67, pp. 275–285, 2018.
- [6] S. Zhang, R. Narayanan, and A. Choudhary, “Voice of the customers: Mining online customer reviews for hotel service improvement,” in *Proc. IEEE Int. Conf. Data Mining Workshops*, 2010, pp. 199–204.
- [7] P. Phillips, S. J. Barnes, K. Zigan, and R. Schegg, “Understanding the impact of online reviews on hotel performance: An empirical analysis,” *J. Travel Res.*, vol. 56, no. 2, pp. 235–249, 2017. doi: 10.1177/0047287516636481.
- [8] B. B. Cao, A.-S. Shia, and Y. Fan, “Study on the impact of online hotel reviews on consumer purchase intention—A case study of S hotel,” *Int. J. Business Manag. Invention*, vol. 13, no. 7, pp. 40–53, 2024.
- [9] N. T. Vo, V. V. Hung, Z. Tučková, N. T. Pham, and L. H. L. Nguyen, “Guest online review: An extraordinary focus on hotel users’ satisfaction, engagement, and loyalty,” *J. Quality Assurance Hospitality Tourism*, vol. 23, no. 4, pp. 913–944, 2021. doi: 10.1080/1528008X.2021.1920550.
- [10] Y. Wen, Y. Liang, and X. Zhu, “Sentiment analysis of hotel online reviews using the BERT model and ERNIE model—Data from China,” *PLOS ONE*, vol. 18, no. 3, e0275382, 2023. doi: 10.1371/journal.pone.0275382.
- [11] M. Sultana and M. M. Rahman, “Service quality of five star hotels in Bangladesh: An empirical assessment,” *Asian Business Rev.*, vol. 3, no. 1, pp. 52–58, 2013.
- [12] D. Chi, T. Huang, Z. Jia, and S. Zhang, “Research on sentiment analysis of hotel review text based on BERT–TCN–BiLSTM–Attention model,” *Array*, vol. 25, art. 100378, 2025.

- [13] K. P.-S. Tan, S. Zou, and X. (R.) Li, “Implications of electronic word-of-mouth on hotel profitability: An investigation beyond room revenue,” *Asian J. Business Res.*, vol. 13, no. 1, pp. 41–63, 2023.
- [14] C. So, “Understanding the prediction mechanism of sentiments by XAI visualization,” in *Proc. 4th Int. Conf. Natural Language Processing and Information Retrieval (NLPAIR)*, 2020.
- [15] M. M. Mariani, I. C. C. Chan, and A. P. Kirilenko, “Sentiment analysis in hospitality and tourism: A thematic and methodological review,” *Int. J. Contemp. Hospitality Manag.*, vol. 34, no. 1, pp. 46–82, 2022.
- [16] A. Shukla and A. Mishra, “Role of review length, review valence and review credibility on consumer’s online hotel booking intention,” *FIIB Business Rev.*, vol. 12, no. 4, pp. 403–414, 2023. doi: 10.1177/23197145221099683.

## ORIGINALITY REPORT

15%

SIMILARITY INDEX

12%

INTERNET SOURCES

9%

PUBLICATIONS

8%


STUDENT PAPERS

## PRIMARY SOURCES

1	Submitted to Midlands State University Student Paper	3%
2	Submitted to Daffodil International University Student Paper	1%
3	S.P. Jani, M. Adam Khan. "Applications of AI in Smart Technologies and Manufacturing", CRC Press, 2025 Publication	1%
4	www.mdpi.com Internet Source	1%
5	Submitted to Universiti Malaysia Pahang Student Paper	<1%
6	annals-csis.org Internet Source	<1%
7	arxiv.org Internet Source	<1%
8	Amin Mojoodi, Tafazal Kumail, Seyed Mohammad Ahmadzadeh, Saeed Jalalian. "Perceptual mapping and key factors influencing hotel choices: A web mining approach to Booking.com reviews", International Journal of Hospitality Management, 2025 Publication	<1%
9	Submitted to Victorian Institute of Technology Student Paper	<1%

10	<a href="http://ebin.pub">ebin.pub</a> Internet Source	<1 %
11	<a href="http://joiv.org">joiv.org</a> Internet Source	<1 %
12	<a href="http://jutif.if.unsoed.ac.id">jutif.if.unsoed.ac.id</a> Internet Source	<1 %
13	<a href="http://repository.itelkom-pwt.ac.id">repository.itelkom-pwt.ac.id</a> Internet Source	<1 %
14	<a href="http://jict.ilmauniversity.edu.pk">jict.ilmauniversity.edu.pk</a> Internet Source	<1 %
15	<a href="http://www.pkm.tunasbangsa.ac.id">www.pkm.tunasbangsa.ac.id</a> Internet Source	<1 %
16	<a href="http://umpir.ump.edu.my">umpir.ump.edu.my</a> Internet Source	<1 %
17	<a href="http://calhoun.nps.edu">calhoun.nps.edu</a> Internet Source	<1 %
18	Submitted to Sunway Education Group Student Paper	<1 %
19	Nazmul Siddique, Mohammad Shamsul Arefin, K. M. Azharul Hasan, M. Shamim Kaiser. "Data Driven Applications for Industry 4.0 and Beyond", CRC Press, 2025 Publication	<1 %
20	Submitted to University of Hertfordshire Student Paper	<1 %
21	<a href="http://iieta.org">iieta.org</a> Internet Source	<1 %
22	"Data Mining and Information Security", Springer Science and Business Media LLC, 2026 Publication	<1 %

# Account Clearance

OMAR FARUK HIMEL  
221-35-846

- Dashboard
- Student Profile
- Payment Ledger
- Registration/Exam Clearance
- Registered Course
- Result
- Routine
- Live Result
- Teaching Evaluation
- Scholarship
- Convocation Apply
- Certificate & Transcript
- Laptop
- Mentor Meeting
- Transport Card Apply
- Student Application
- Logout

### Dashboard

Student Portal


Total Payable	Total Paid	Total Due	Total Other
767,200.00	767,200.00	0.00	1,500.00

Today's Routine - Saturday

No routine available for today.

### Semester Wise Result

Semester-wise SGPA Performance



Semester	SGPA
Spring, 2022	3.48
Summer, 2022	3.25
Fall, 2022	3.88
Spring, 2023	3.98
Fall, 2023	3.98
Spring, 2024	3.34
Fall, 2024	3.34
Spring, 2025	3.31
Summer, 2025	2.96

2025 © - Developed by Daffodil International University