5-11-2024

A Framework for Filtering Irrelevant NOTAMs

Roudha Abdulrahman
raa1759@rit.edu

A Framework for Filtering Irrelevant NOTAMs

by

Roudha Abdulrahman

A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Master of Science in Professional Studies: Data Analytics

Department of Graduate Programs & Research

Rochester Institute of Technology
RIT Dubai
May 11, 2024

RIT

Master of Science in Professional Studies:
Data Analytics

Graduate Thesis Approval

Student Name**: Roudha Abdulrahman**
Graduate Thesis Title**: A Framework for Filtering Irrelevant NOTAMs**

Graduate Thesis Committee:

Name:	**Dr. Sanjay Modak** **Chair of committee**	**Date:**
Name:	**Dr. Ehsan Warriach** **Mentor**	**Date:**

Acknowledgments

I would like to express my deepest appreciation to Rochester institute of Technology for providing the facilities and resources needed to conduct this study.

For their helpful advice, encouragement, and support during this research journey, I am incredibly grateful to my mentor Dr. Ehsan Warriach. His knowledge, tolerance, and perceptive criticism have been invaluable in helping to shape this thesis.

I am grateful to my family for their support and understanding; they have been a continual source of inspiration and strength.

I want to express my gratitude to everyone who took part in the study and to those who gave their time and wisdom.

Lastly, I would like to express my gratitude to all of my friends and coworkers who have encouraged, understood, and been a friend to me along this journey.

The combined efforts and assistance of everyone listed above made this work possible.

Abstract

The Notice to Airmen (NOTAM) system is essential for aviation safety, giving critical information on risks and operating limitations. However, the volume and complexity of NOTAM data complicate interpretation, which could compromise safety and efficiency. This study tries to address these issues by creating a predictive algorithm for analysing NOTAM data and predicting whether to keep or remove them. The methodology combines effective machine learning algorithms to reveal insights that improve safety and operational efficiency. Using a dataset of NOTAM entries, multiple prediction algorithms are tested to forecast possible problems and enable proactive risk management. The study's findings help to improve the effectiveness of NOTAM prediction systems, resulting in increased aviation safety.

The proposed model for this study is Support Vector Machine model with the TF-IDF transformed data, the classification performance was assessed using standard evaluation measures such as accuracy, precision, recall, F1-score and ROC. The results showed that the SVM model was 76% accurate in categorizing NOTAM entries as "keep" or "remove."

Keywords: Notices to Airmen (NOTAMs), aviation safety, automated classification, machine learning, natural language processing (NLP), supervised learning, unsupervised learning, data science, operational efficiency.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Introduction

This chapter gives a summary of the research topic, including background information, its importance of the study, and the problem statement. It also introduces the research questions and objectives, as well as the thesis structure.

1.2 Background Information

NOTAMs (Notices to Airmen) are critical pieces of information communicated by civil aviation authorities, serving as vital communication tools within the aviation industry. These notices are designed to inform and alert pilots to various modifications, updates, or temporary changes in the airspace and related infrastructure, such as changes in airport operations, changes in flight routes, amendments to local procedures, and the presence of potential hazards.

NOTAMs are crucial because they are essential to maintaining both operational effectiveness and aircraft safety. It is as well part of the Operational Flight Plan (OFP), which includes all flight-related information and instructions. OFPs are detailed documents that include information such as the flight route, weather predictions, fuel requirements, and any pertinent notices or limitations affecting the trip. These alerts help prevent possible dangers, provide a proactive reaction to new difficulties, and maintain a secure operating environment by giving vital information to flight crews, ground staff, and airspace management. NOTAMs serve as a channel for the dissemination of urgent, information that pilots and aviation stakeholders need to know to make educated judgements and modify their operational plans in response to changing conditions or changes in the airspace environment. They are essential for reducing hazards, improving safety procedures, and facilitating smooth air traffic movement. Thus, ensuring the safety, effectiveness, and compliance of aviation operations requires a thorough understanding of and application of NOTAMs.

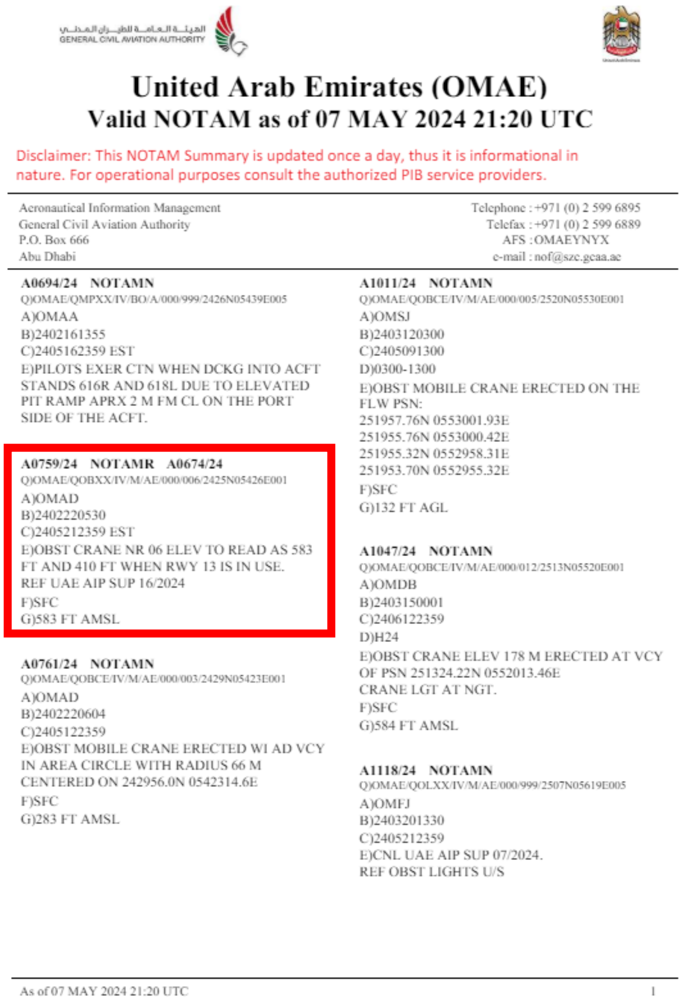
Example of a NOTAM:

A0759/24 A0674/24
A)OMAD
B)2402220530
C)2405212359 EST
E)OBST CRANE NR 06 ELEV TO READ AS 583 FT AND 410 FT WHEN RWY 13 IS IN USE. REF UAE AIP SUP 16/2024 NOTAMR
Q)OMAE/QOBXX/IV/M/AE/000/006/2425N05426E001
F)SFC
G)583 FT AMSL

1. **Unique Identifier:** A0759/24 and A0674/24
2. **Location:** A) OMAD indicates that the NOTAM is for Abu Dhabi International Airport (OMAD).
3. **Nature of the NOTAM:** E) OBST CRANE indicates that the NOTAM is related to an obstacle, specifically a crane.

4. **Details of the Obstacle:** The crane's elevation is specified as 583 feet above mean sea level (FT AMSL) and 410 feet when runway 13 is in use.
5. **Reference:** REF UAE AIP SUP 16/2024 refers to the United Arab Emirates (UAE) Aeronautical Information Publication (AIP) Supplement 16/2024.
6. **Status:** NOTAMR indicates that this is a current NOTAM.
7. **Category:** Q) OMAE/QOBXX/IV/M/AE/000/006/2425N05426E001 specifies the category of the NOTAM.
8. **Start and End Times:** B)2402220530 specifies the start time of February 24th, 2022, at 0530 UTC, and C)2405212359 EST specifies the end time of May 24th, 2021, at 2359 UTC.
9. **Affected Area:** F) SFC indicates that the surface area is affected.
10. **Additional Information:** G) 583 FT AMSL provides additional information about the elevation of the obstacle.

NOTAM description: This NOTAM notifies of an obstacle crane near runway 06, with an elevation of 583 feet and 410 feet when runway 13 is operational. It pertains to the UAE Aeronautical Information Publication Supplement 16/2024. Active from February 24th, 2022, at 0530 UTC until May 24th, 2021, at 2359 UTC, affecting surface area.


GENERAL CIVIL AVIATION AUTHORITY

United Arab Emirates (OMAE)
Valid NOTAM as of 07 MAY 2024 21:20 UTC

Disclaimer: This NOTAM Summary is updated once a day, thus it is informational in nature. For operational purposes consult the authorized PIB service providers.

Aeronautical Information Management
General Civil Aviation Authority
P.O. Box 666
Abu Dhabi

Telephone : +971 (0) 2 599 6895
Telefax : +971 (0) 2 599 6889
AFS : OMAEYNYX
e-mail : nof@szc.gcaa.ae

A0694/24 NOTAMN
Q)OMAE/QMPXX/IV/BO/A/000/999/2426N05439E005
A)OMAA
B)2402161355
C)2405162359 EST
E)PILOTS EXER CTN WHEN DCKG INTO ACFT STANDS 616R AND 618L DUE TO ELEVATED PIT RAMP APRX 2 M FM CL ON THE PORT SIDE OF THE ACFT.

A0759/24 NOTAMR A0674/24
Q)OMAE/QOBXX/IV/M/AE/000/006/2425N05426E001
A)OMAD
B)2402220530
C)2405212359 EST
E)OBST CRANE NR 06 ELEV TO READ AS 583 FT AND 410 FT WHEN RWY 13 IS IN USE. REF UAE AIP SUP 16/2024
F)SFC
G)583 FT AMSL

A0761/24 NOTAMN
Q)OMAE/QOBCE/IV/M/AE/000/003/2429N05423E001
A)OMAD
B)2402220604
C)2405122359
E)OBST MOBILE CRANE ERECTED WI AD VCY IN AREA CIRCLE WITH RADIUS 66 M CENTERED ON 242956.0N 0542314.6E
F)SFC
G)283 FT AMSL

A1011/24 NOTAMN
Q)OMAE/QOBCE/IV/M/AE/000/005/2520N05530E001
A)OMSJ
B)2403120300
C)2405091300
D)0300-1300
E)OBST MOBILE CRANE ERECTED ON THE FLW PSN:
251957.76N 0553001.93E
251955.76N 0553000.42E
251955.32N 0552958.31E
251953.70N 0552955.32E
F)SFC
G)132 FT AGL

A1047/24 NOTAMN
Q)OMAE/QOBCE/IV/M/AE/000/012/2513N05520E001
A)OMDB
B)2403150001
C)2406122359
D)H24
E)OBST CRANE ELEV 178 M ERECTED AT VCY OF PSN 251324.22N 0552013.46E
CRANE LGT AT NGT.
F)SFC
G)584 FT AMSL

A1118/24 NOTAMN
Q)OMAE/QOLXX/IV/M/AE/000/999/2507N05619E005
A)OMFJ
B)2403201330
C)2405212359
E)CNL UAE AIP SUP 07/2024.
REF OBST LIGHTS U/S

As of 07 MAY 2024 21:20 UTC 1

Figure 1.1: Sample of NOTAMs.

1.3 Project Goals

The main objective of this project is to use machine learning techniques to create an effective framework for filtering out irrelevant Notices to Airmen (NOTAMs) in the aviation industry. The model is intended to provide accurate prediction results to assist in categorizing NOTAMs as 'keep' or 'remove'.

1.4 Aims and Objectives

1. Collect and preprocess NOTAM data to produce a comprehensive dataset suitable for training machine learning models.
2. Train machine learning models that can correctly distinguish between relevant and irrelevant NOTAMs based on their content and context.
3. Perform testing and validation processes to determine the performance and effectiveness of the generated models in filtering out irrelevant NOTAMs.
4. Reduce fatigue on pilots through reducing the read of long OFP's and going through each NOTAM relevant and irrelevant ones. They don't miss any of the critical NOTAMs.
5. Ensure pilots do not miss any critical NOTAMs, hence improve safety.

6. Work with industry stakeholders to incorporate the developed machine learning framework into operational workflows and deploy it in real-world aviation environments.

1.5 Research Methodology

The study leverages the Notice to Airmen (NOTAM) data to improve aviation and operational efficiency.

Dataiku, an advanced data science platform will be utilized to complete this project. Dataiku has excellent features for data preparation, analysis, and machine learning, making it a suitable platform for the NOTAM data investigation. Using the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology, the project proceeds through the following key phases:

1.5.1 Business understanding

The aviation industry relies significantly on the Notice to Airmen (NOTAM) system as a key communication tool. NOTAMs provide critical information about potential risks, operational changes, and airspace restrictions at airports to ensure the safety and efficiency of air travel. However, the huge number of NOTAMs created every day presents a major challenge to aviation industry. Among this large volume of information, identifying important and irrelevant NOTAMs becomes increasingly challenging, resulting in inefficiencies in information processing and significant safety issues.

The advent of technology has ushered in innovative ways to the difficulty of managing Notice to Airmen (NOTAM) data in the aviation sector. Machine learning techniques have enabled the development of sophisticated algorithms capable of effectively filtering out irrelevant NOTAMs. These machine learning models can assess large amounts of NOTAM data, learning patterns, and contexts to accurately identify and prioritize relevant information.

1.5.2 Data understanding

The dataset for this study is made up of Notice to Airmen (NOTAM) texts sourced from aviation database and archives. Each NOTAM entry contains information on aircraft operations, such as runway closures, airspace restrictions, navigational assistance outages, and other critical updates. The collection comprises a wide range of NOTAMs issued by aviation authorities, providing a comprehensive picture of airspace management and flight safety alerts.

1.5.3 Data preparation

The data preparation process will encompass collection, cleaning, preprocessing, and text-handling steps to ensure the dataset's readiness for analysis. By systematically preparing the dataset, we aim to facilitate the application of machine learning techniques for classifying the relevance of NOTAMs and aiding in aviation decision-making processes.

1.5.4 Modelling

The use of machine learning techniques has increased dramatically in recent years, owing to their ability to improve data analysis and decision-making processes. These algorithms, known for their scalability and efficacy in handling large datasets. Machine learning algorithms, particularly in the context of Notice to Airmen (NOTAM) data, provide the ability to discover nuanced correlations between textual elements that are difficult for humans to perceive. By exploiting these correlations, these models can make precise predictions or classifications, assisting in determining NOTAM relevance and significance for aircraft operations.

In this study, two machine learning classifiers (Random Forest and Support Vector Machine) will be trained on extracted features derived from various text handling techniques, such as TF-IDF vectorization and sentence embeddings using DistilBERT and MPNet, to classify NOTAMs accurately and efficiently, allowing for more informed decision-making in aviation management. This strategy was chosen based on successful practices and previous research, which will be thoroughly reviewed in the literature review section.

1.5.4.1 Random Forest

The random forest technique is an extension of the bagging approach by combining bagging and feature randomness to generate an uncorrelated forest of decision trees. Feature randomness, also known as feature bagging or "the random subspace method" , provides a random subset of features, ensuring low correlation among decision trees. This is an important distinction between decision trees and random forests. Random forests select only a subset of the available feature splits, whereas decision trees consider all of them.

Random forest techniques include three primary hyperparameters that must be configured before training. These include node size, number of trees , and number of features sampled. From there, you can utilize the random forest classifier to tackle regression or classification problems.

The random forest algorithm consists of a collection of decision trees, with each tree in the ensemble containing a data sample selected from a training set with replacement, known as the bootstrap sample. One-third of the training sample is set aside as test data, also known as the out-of-bag (oob) sample. . Another instance of randomness is then injected by feature bagging, increasing the dataset's variety and decreasing correlation among decision trees. Depending on the type of problem, the prediction will differ. In a regression job task the individual decision trees will be averaged, but in a classification task, the predicted class will be determined by a majority vote—that is, the most common categorical variable. Finally, the oob sample is used for cross-validation, which finalizes the prediction.

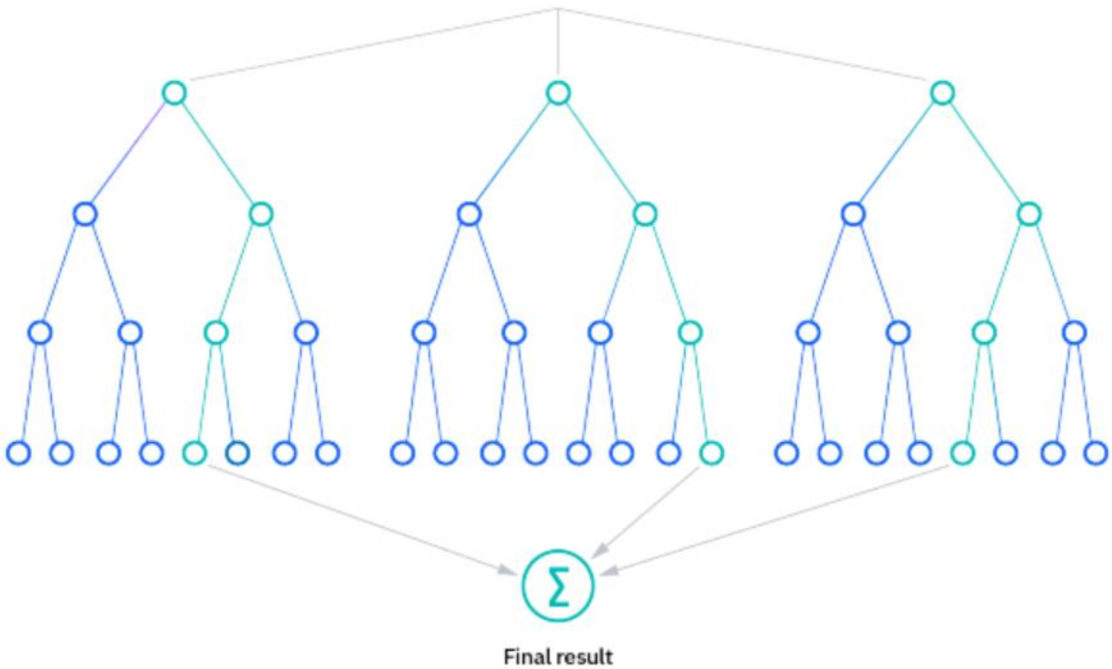


Figure 1.2: Random Forest Representation.

1.5.4.2 Support Vector Machine

The Support Vector Machine (SVM) works on the concept of defining an optimal hyperplane that maximally separates data points from distinct classes. In the domain of NOTAM classification, SVM seeks to learn a decision boundary that can effectively distinguish between relevant and irrelevant notices.

Mathematically, SVM creates this decision boundary by identifying the best hyperplane that maximizes the margin between the nearest data points of various classes, also known as support vectors.

The selection of an appropriate kernel function is critical in SVM since it dictates how input data is transformed into feature space. Kernel functions such as linear, polynomial, radial basis function (RBF), and sigmoid are commonly utilized, with each providing varying degrees of flexibility and expressiveness in capturing complicated relationships within data.

1.5.5 Evaluation

The models will be evaluated using the following evaluation metrics:

1.5.5.1 Accuracy

Accuracy is a basic metric for assessing classification models. It represents the proportion of observations correctly predicted by the model. It is calculated using the following formula:

$$accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

1.5.5.4 Precision

Precision is the proportion of positive outcomes that have been correctly classified. It is calculated using the following formula:

$$precision\ = \frac{TP}{TP + FP}$$

1.5.5.4 Recall

Recall is the percentage of true positive classifications (True Positives) from actual positive cases.

$$precision\ = \frac{TP}{TP + FP}$$

1.5.5.5 F1-score

The F1 score refers to the harmonic mean between precision and recall.

$$Precision\ = \frac{Precision *\ Recall}{Precision +\ Recall}$$

1.5.5.6 ROC

The Receiver Operating Characteristic (or ROC) curve shows the true positive rate vs. the false positive resulting from different cutoffs in the predictive model. The steeper the curve, the better. On the contrary, a curve close to the diagonal line is worse.

1.6 Limitations of the Study

1.6.1 Feature Selection

The Feature selected for model training may have had an impact on the framework's performance and efficacy. Alternative feature sets or extraction approaches may produce different outcomes and should be investigated in future research.

1.6.2 Abbreviation Expansion

The inclusion of NOTAMs having abbreviations that needed to be expanded presented a hurdle in this investigation. Several limitations were discovered while attempting to address this issue by mapping abbreviation translations with a dictionary. For example, some abbreviations may have several translations, resulting in ambiguity when expanded. Despite efforts to overcome these obstacles, remaining concerns with abbreviation expansion may have impacted the dataset's correctness and completeness.

1.6.3 Training data size

The small size of the NOTAM training dataset presents a substantial barrier, impeding the creation of accurate and generalizable models. The lack of labelled NOTAM data limits the ability of machine learning algorithms to generalize successfully, potentially leading to overfitting and poor performance on unlabelled data.

Chapter 2: Literature Review

2.1 Introduction

This chapter provides a critical review of previous studies and research on NOTAMs and the application of AI and machine learning for prediction. It emphasizes the strengths and limits of prior research and identifies knowledge gaps that the thesis aims to fill.

2.2 Literature Review

Notices to Airmen, or NOTAMs, are essentially messages or notification sent to aviation personnel. These notifications, which help pilots, ground crews, and other personnel understand dynamic changes that might affect their flights, are crucial for flight safety as they convey abnormal or changing statuses within the National Airspace System. They play a vital role in distributing valuable information about weather, hazards, and no-fly zones, encouraging decision-makers to make safe and informed decisions about flying. Furthermore, the International Civil Aviation Organization (ICAO) format is emphasized as a step toward global uniformity and improved access to accurate information, replacing domestic NOTAMs. NOTAMs come in a variety of forms to address different flight scenarios. Trigger NOTAMs, Temporary Flight Restrictions (TFRs), Flight Data Center (FDC) notifications, ASHTAMs, SNOWTAMs, BIRDTAMs, and the ICAO format inform people about things like ash, snow, and bird hazards. It is recommended that pilots, ground crews, and other aviation-related personnel include the most recent NOTAMs in their flight kits to facilitate convenient access to vital operational data.

The adoption of telecommunications in 1947 marked the beginning of NOTAM issuance which was inspired by the Notice to Mariners used to notify ship captains of perils at sea. NOTAMs, which are intended for efficient communication, use a different language with specialized contractions. These contractions are crucial in improving communication efficiency and assisting computer systems in digesting critical information. The International Civil Aviation Organization (ICAO) is the principal authority in the global aviation community for standardizing contractions. When ICAO contractions are not available, plain language is used. The NOTAM is constructed in a specified order, with accountability coming first, followed by the NOTAM number, impacted location, keyword, and the components of Start of Activity and End of Validity always coming last.

A study recently analyzed the impact of NOTAMs on flight safety and efficiency, with a focus on major concerns such as the massive number of NOTAMs, practical usage challenges, and technical limitations in the current system. NOTAM's negative effects on air navigation support are demonstrated, as are worldwide best practices such as the use of Q-codes and advanced flight planning systems such as Lido Flight 4D. The study investigates topics such as the European AIS Database (EAD), which is based on the Aeronautical Information interchange Model (AIXM), and Digital NOTAM, which is utilized for dynamic aeronautical data interchange. The investigation focuses on the FAA's improved Federal NOTAM System (FNS), which can encode Digital NOTAMs.

In order to ensure accurate and dependable flight plans, Lido Flight 4D uses a combination of automation and human intervention for the interpretation of Notice to Airmen (NOTAMs). Around 70% of NOTAMs are processed automatically by the system because they deal with simple, repetitive cases. However, because of their varied complexity, subtle linguistic nuances, and contextual significance, the remaining notifications require human analysis. To ensure that the system applies NOTAMs to flight planning actions for all Lido customers, the human task entails translating and clarifying restrictions. In order to ensure accurate implementation within the flight planning system, human involvement is crucial when handling NOTAMs that could affect various aspects, such as runway closures, equipment failures, or restricted airspace.

The existing NOTAMs system is regarded as "clumsy" and error-prone, necessitating the need for intelligent querying and filtering. The move to Aeronautical Information Management (AIM) has resulted in the introduction of Digital NOTAMs, and the Semantic NOTAM-system (SemNOTAM) solves these problems by utilizing a knowledge representation and reasoning system. The system helps with airspace efficiency and trajectory control by supporting SESAR, AIRM, and ISRM aims. Personalization, user profiles, and incorporation into electronic flight baggage for in-flight assistance are among the topics that will be covered in future work. SemNOTAM is a promising development in NOTAM handling that overcomes existing constraints and fits in with the changing aeronautical information management environment. SemNOTAM makes use of Digital NOTAMs to allow for fine-grained filtering depending on event type, location, aircraft, time, and user preferences. It isolates data representation from business rules, improving interoperability and making semantic descriptions easier. SemNOTAM, in contrast to previous systems, enables intelligent filtering, decreasing information overload, allowing user-defined prioritizing, and harmonizing with AIM's data-centric orientation.

One of the most important aspects of contemporary air travel is the improvement of aviation safety through excellent communication. In order to effectively categorize aviation warning communications, or "Notice To Airmen" (NOTAM), using artificial intelligence (AI), Jean Coupon, an astronomer and former SIT Academy data science student, worked with SWISS International Airlines. A lower safety threshold frequently causes an increase in NOTAM volume, which increases the possibility that pilots would overlook important information. The research developed an automatic NOTAM classifier using Natural Language Processing (NLP) and Machine Learning. The procedure comprised labelling, grouping related NOTAMs, and translating individual words into computer representations. With just one month's worth of data, the supervised Machine Learning algorithm was able to train a model with an astounding 94% accuracy in determining the significance of NOTAMs. In order to continuously enhance the system, the ongoing research intends to evaluate more data, achieve near-perfect accuracy for particular message clusters, and include expert feedback. This creative method shows how artificial intelligence (AI) may improve airline safety and streamline communication procedures.

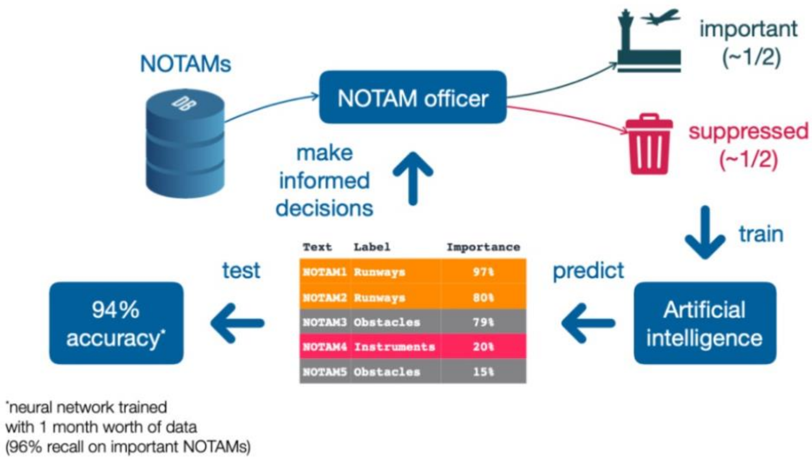


Figure 2.1: NOTAM categorization Solution by Jean Coupon.

A study by Banane and Erraissi (2020) examined how big data technologies and natural language processing (NLP) interact, with a particular emphasis on how the latter improves the administration of large-scale NLP datasets. The goal of natural language processing (NLP), a vital area of artificial intelligence, is to comprehend spoken and written human language. Specialised computer programmes that use syntactic and semantic analyses to determine sentence meaning based on predefined grammatical rules make this comprehension easier. By comparing text with real-time databases, these programmes may split words and groups of words or analyse the grammar of entire sentences to determine meaning and context. Effective NLP algorithms, however, rely significantly on having access to vast amounts of data in order to find relevant relationships. Simultaneously, Big Data technologies transform the way Big Data are managed through the use of dynamic systems like NoSQL and Spark. A comparative analysis emphasising the benefits of using Big Data for managing large NLP datasets follows Banane and Erraissi's thorough investigation of NLP strategies utilising Big Data technologies.

A publication featured in the Collegiate Aviation Review (2022) delve into the significant impact of Natural Language Processing (NLP) on the aviation sector. They pinpoint three crucial domains where NLP is making strides: safety report analysis, aviation maintenance, and air traffic control. By leveraging NLP, stakeholders can classify safety reports effectively, predict maintenance issues, and enhance communication protocols between controllers and pilots. The authors underline the increasing integration of NLP software across these sectors, emphasizing its potential to revolutionize traditional practices and improve overall operational efficiency within the aviation industry. However, they also acknowledge the existing challenges and propose avenues for future research to further optimize NLP applications in aviation.
Another study presented a novel approach that makes use of Natural Language Processing (NLP) algorithms in order to address the growing difficulties that the human processing of the increasing number of Notices to Airmen (NOTAMs) in the international civil aviation industry presents. The research applies a complex NOTAM Information Extraction (NIE) model with Bi-

LSTM+CRF and describes a hierarchical framework, the Classification and Hierarchical Structure for NOTAM Information (CHNI). By methodically arranging NOTAM data in a hierarchical fashion, CHNI establishes the foundation for the NIE model. Testing the NIE model on the airplane navigation information system of the Chinese Mainland reveals that it is remarkably accurate—it surpasses 92% for both two- and three-level events. This innovation demonstrates how NLP technology may greatly improve NOTAM processing efficiency in scenarios including navigational information for civil aviation.

As mentioned earlier, during pre-flight briefings, pilots must analyze specialist NOTAMs, which are written in a distinct language with acronyms and domain-specific words. Using self-supervised language models like BERT that have been pre-trained on a sizable collection of NOTAMs, a novel method addresses this. These models support named entity recognition, translation for pilot applications, and criticality prediction. For fine-tuning, the approach effectively uses less labeled data. The downstream activities are designed to assist pilots in rapidly determining the significance of NOTAMs and extracting pertinent data. It is emphasized that these models have the potential to be operationally useful and that they might greatly benefit aeronautical applications. To ensure reliability and certification in aviation applications, future research could investigate alternative natural language processing (NLP) methods and address concerns regarding model overconfidence in safety-critical scenarios. The study essentially emphasizes the use of cutting-edge natural language processing (NLP) methods to improve NOTAM analysis for better pilot decision-making and operational efficiency.

This NASA-published study used machine learning techniques to evaluate Notices To Airmen (NOTAMs), yielding noteworthy findings. Using TF-IDF embeddings and k-means clustering, the study successfully classified NOTAMs into groups such as Airport, Weather, and Airspace. Furthermore, Latent Dirichlet Allocation (LDA) uncovered hidden subjects, improving subject matter predictions and NOTAM similarity. A Named Entity Recognition (NER) model trained on annotated NOTAMs got an astounding 98% F1-score, revealing structural differences. The usefulness of RoBERTa and XLNet for question answering with plain language NOTAMs was demonstrated. The goal of future research is to create a mechanism for translating handwritten NOTAMs to digital format. Furthermore, the research team intends to apply Natural Language Processing (NLP) approaches to additional air traffic management documents, such as Letters of Agreement (LoAs) and Standard Operating Procedures (SOPs), in order to inform National Airspace System (NAS) operations. Ongoing research is focusing on new NLP tasks, such as optimizing question answering for broader air traffic management applications and incorporating hybrid models for improved data quality and user interaction within standardized aviation models such as AIXM.

Preprocessing techniques are critical in the conversion process because they ensure that textual data is properly cleaned and filtered, which improves the effectiveness of later information retrieval strategies. A survey done by Tabassum and Patil (2020) in the International Research Journal of Engineering and Technology (IRJET) emphasises the importance of text preprocessing and feature extraction approaches in NLP. This survey emphasises the necessity of using techniques like tokenization, stop word removal, punctuation removal, and lemmatization to increase the accuracy of text-based machine learning algorithms. It also lists popular feature extraction approaches such as the Bag-of-words model and TF-IDF, with TF-IDF being especially useful in web search or search engine applications."

Transformers are a deep learning architecture that has recently transformed the field of natural language processing (NLP). They are commonly used for tasks like as language translation, text classification, sentiment analysis, and so on. The transformer design consists of an encoder and a decoder, each with multiple layers of self-attention and feedforward neural networks. The transformer's heart is the self-attention mechanism, which allows the model to prioritize distinct words in a sentence based on their affinity with one another. This is comparable to how a human would read a statement, focusing on the most important sections rather than reading it in order from beginning to end. In addition to self-attention, the transformer incorporates positional bias, which enables the model to track the relative placements of words in a sentence. This is crucial because the arrangement of words in a phrase can considerably influence its meaning.

The transformer encoder architecture is used in applications such as text classification, where it processes a sequence of tokens to create a fixed-size vector representation that is utilized for classification. BERT is a well-known model in this area, because to its capacity to pre-train on large amounts of text data and adapt to a variety of NLP applications.

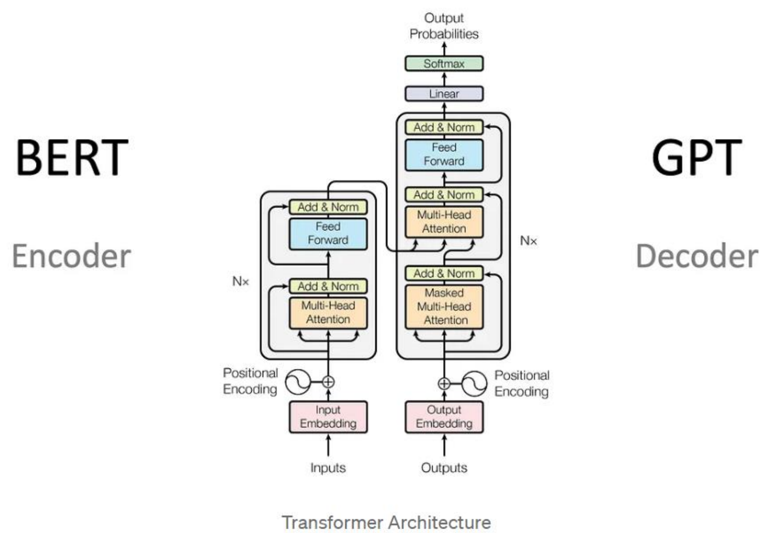


Figure 2.2: Transformer Architecture

In contrast, the transformer decoder architecture is used for tasks like as language synthesis, which involves producing word sequences based on input context. It gets a fixed-size vector representation and generates words sequentially based on previous outputs.

Google's BERT and subsequent transformer-based models have become industry standard in natural language processing (NLP) because of their outstanding performance on a wide range of

jobs. More recently, efforts have focused on improving computing efficiency or forecast accuracy, but not both at the same time.
By using permutation language modeling, XLNet outperforms BERT on a range of language tasks, resulting in improved performance. Developed at Facebook, RoBERTa outperforms BERT and XLNet with robust optimization achieved through improved training methods and a larger dataset.

DistilBERT provides a simplified substitute by condensing the knowledge of BERT into a smaller network, enabling faster inference without a discernible reduction in prediction accuracy.

Which of these models to choose will depend on your particular demands. DistilBERT is the best choice for applications that prioritize faster inference, while RoBERTa is the better choice for workloads that require optimal prediction metrics. Permutation-based training of XLNet, on the other hand, appears promising for managing intricate dependencies. Overall, these developments highlight the continued development of NLP and offer useful choices for scholars and practitioners.
Text classification is a key task in the field of Natural Language Processing (NLP), with several applications, including sentiment analysis and topic modelling. Significant progress has been made in text classification with the introduction of machine learning methods. Nonetheless, worries over data privacy have led academics to investigate ways to classify text on encrypted data. The combination of machine learning and cryptography in text categorization is examined in this review of the literature, with a particular emphasis on a paper published and presented at the 2023 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). The study explores the use of different machine learning algorithms on encrypted text data to show that it is possible to protect classification accuracy and data confidentiality at the same time.
In an article authored by Ruben Winastwan's (2024), he focuses on improving Natural Language Processing (NLP) with AI approaches, namely sentence embedding. The drawbacks of traditional embedding approaches are addressed, and Sentence Transformers are proposed as a solution that successfully leverages the architecture of Transformer models. The essay delves into MPNet, a paradigm that combines the benefits of BERT and XLNet while resolving their limitations. It shows how MPNet combines Masked Language Modelling (MLM) and Permutation Language Modelling (PLM), with two-stream self-attention for efficient contextual embeddings. In addition, insights into MPNet implementation and fine-tuning are offered, with examples of its applicability in various NLP tasks such as text categorization and document summarization.
Support Vector Machine (SVM) is a machine learning technique for classification and regression that specializes at binary classification. It aims to determine the best decision boundary, or hyperplane, for separating various classes by maximizing the margin between data points. SVMs may handle both linearly and non-linearly separable data by converting it to a higher-dimensional space using kernel functions. While SVMs have advantages such as resistance to overfitting and adaptability in handling high-dimensional data, they require parameter tuning and may encounter computing difficulties with large datasets. Nonetheless, SVMs remain an effective tool in a variety of machine learning applications.

A Study by Liu et al. (2010) focuses on SVM Compared with the other Text Classification Methodsuse of Support Vector Machines (SVMs) in text categorization and demonstrates its efficacy in improving classification results. The article discusses the fundamental principles of SVM and describes the text categorization process, before providing an SVM-based model. Experimental results show that the SVM classifier obtains an F1 value greater than 86.26%, showing a substantial improvement over conventional classification method. The review acknowledges the problems of SVM, such as the subjective selection of error parameters (C and γ) and the kernel function. Despite its efficiency, SVM requires additional theoretical and practical advances to address these issues and improve its use in text categorization.
Ohata et al. (2022) conducted research on developing an assistant for a large company's technical support system, which handles asynchronous services for resolving technical problems. The goal is to use natural language processing (NLP) and machine learning algorithms to suggest appropriate response templates based on received issue descriptions. The proposed modular pipeline preprocesses raw text input, extracts features, and uses supervised learning to offer appropriate templates or flag issues that require manual resolution. In real-world testing, the pipeline produced an impressive average accuracy of 72.7% across nine classes despite having minimal labelled training data. Furthermore, post-hoc research demonstrated the pipeline's capacity to reliably detect phrases that are closely related to relevant templates. The suggested approach employs TF-IDF encoding, extra information, and a Random Forest classifier, yielding an average accuracy of 72.7% and a weighted F1-score of 69.2%. A qualitative evaluation with the LIME package showed the model's ability to appropriately recognize business-related words. Future work could involve improving the solution with more metadata, conducting semi-supervised learning with unlabelled data, and adopting a sequential learning process.

2.3 Key Takeaways

The evaluation of the existing literature revealed that.

- As the number of international flights increases, airlines prioritize safety by issuing Notices to Air mission (NOTAMs), resulting in an increase in human processing issues.

- Increase in NOTAM volume, probably will increase the possibility that pilots would overlook important information, hence affecting safety.

- NLP and AI approaches like Random Forest and SVM provide ways to automate NOTAM processing, increasing productivity for aviation personnel.

- Text Handling Techniques such as TF-IDF, sentence embedding would be useful for text representations prior to the modelling stage.

- SWISS International Airlines represented as a similar case study, giving a one-month dataset for the NOTAM classifier's initial testing. The success demonstrated in this restricted trial paves the way for future bigger evaluations and deployments across varied datasets.

- Using historical operational data annotated by aviation information professionals, the study validates proposed models, CHNI and NIE, indicating promising outcomes.

Chapter 3: Project Description

3.1 Data Description

3.1.1 Data Sources

The issuance of NOTAMs is a crucial process initiated by aviation authorities to share essential data that may affect flight safety or operations. To ensure accurate representation in pilot briefings and operational flight plans (OFPs), interaction with Lido, a flight planning and navigation program, is required. Predefined filters are used in Lido to effectively sort and prioritize material. Furthermore, the procedure entails creating OFPs that integrate the most recent NOTAM data for pilot use. SharePoint is used to promote collaboration and distribution.

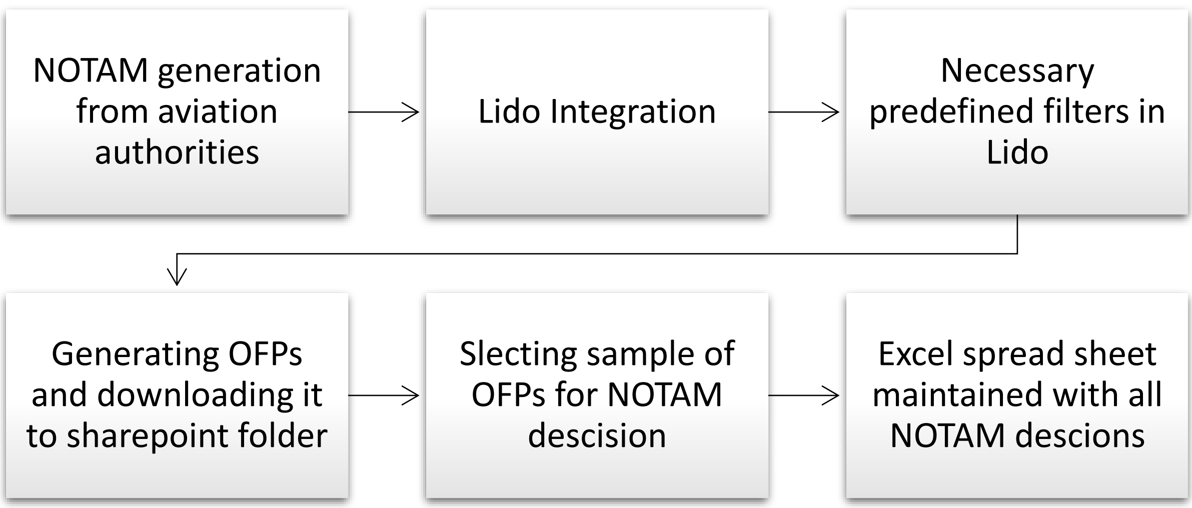


Figure 3.1: Data Source

3.1.2 Data Features

The dataset contains two attributes which are Details and details expanded. The target variable is decision column.

Attribute	Description	Type
Decision (Target Variable)	Final Decision either ‘keep’ or remove.	String
Details	NOTAM text	String

Table 3.1: Attributes, description, and type.

3.1.3 Data Characteristics

The analysis of text length distribution provides useful information about the composition of textual data in the dataset. Descriptive statistics were used to investigate the distributional features of text lengths. The dataset contains 7956 text samples, with an average text length of about 304 characters and a standard deviation of 636. Quartile analysis shows that 25% of the text samples are 78 characters or fewer, 50% are 149 characters or less, and 75% are 309 characters or less. These statistical summaries provide a complete representation of the variation and central trend seen in text lengths across the sample.

Statistic	***Value***
Count	7956
Mean	304
Std	635
25%	78
50%	149
75%	309

Table 3.2: Text length analysis

To identify common language patterns, the frequency distribution of words in the text data was examined. To enable a concentrated examination, common connecting terms such as 'and', 'or', and 'the' were purposefully eliminated. The resulting bar chart depicts the frequency of the top 20 most common words, providing information on dominant phrases and recurring themes in the textual corpus. This visualization helps identify key language elements and content themes by providing a deeper view of the lexical environment found in the dataset.

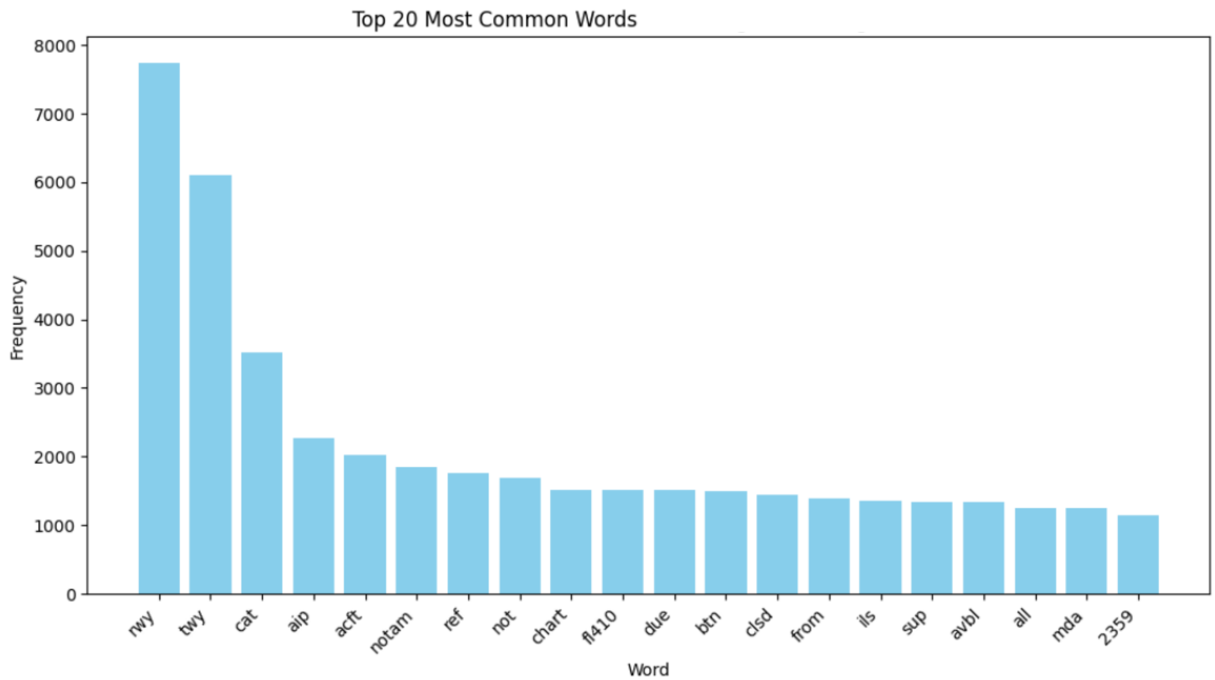


Figure 3.2: Top 20 most common words

The vocabulary size is an important statistic for analyzing the lexical richness and diversity of textual data. In this analysis, vocabulary size was calculated as the total number of unique terms found in all text samples in the dataset.

The calculation yielded a vocabulary size of 29,652 unique words, demonstrating the richness and diversity of lexical elements included in the dataset. This metric emphasizes the complexity and diversity of the language employed in the corpus, giving fundamental insights for future linguistic research and interpretation.

N-grams analysis provides information about the co-occurrence patterns of terms in textual data. Notable linguistic patterns and repeated phrases can be detected by examining adjacent word sequences known as N-grams.

For this research, the most common bi-grams (2-grams) found in the dataset were calculated. The top ten most frequent bi-grams, along with their corresponding frequencies, are listed below:

Rank	**Bi-gram**	**Frequency**
1	('CHART', 'NOTAM')	1308
2	('WILL', 'BE')	1244
3	('OF', 'TWY')	1000
4	('DUE', 'TO')	945
5	('AIP', 'SUPPLEMENT')	851
6	('AND', 'TWY')	821
7	('CAT', 'D:')	805
8	('CAT', 'C:')	804

9	('REF', 'AIP')	788
10	('OF', 'RWY')	637

Table 3.3: Big-grams and frequencies

3.2 Data Preprocessing

3.2.1 Data Cleaning

The target column, which denoted the decision to "keep" or "remove" each NOTAM, contained inconsistencies due to capitalization variances, spelling errors, and duplicate entries. To ensure uniformity and facilitate effective classification, the following cleaning methods were done.

3.2.2 Handling Abbreviations

Abbreviations occurred in 'details' column were expanded in a systematic way to guarantee consistency and clarity across the dataset. A dictionary-based strategy was used, based on the International Civil Aviation Organization's (ICAO) abbreviation and code catalogue. This collection provided a thorough reference for common aviation abbreviations, making it easier to map abbreviations to their complete forms. Additional column 'details_expanded' was added to the dataset.

3.3 Feature Engineering

3.3.1 Feature Selection

Given the nature of the dataset and the modeling methodologies used, explicit feature selection was not conducted. Instead, different text handling strategies were investigated, and the resulting representations of the NOTAM text were used as input features to train both Support Vector Machine (SVM) and Random Forest classifiers.

Text Handling Techniques:

<u>TF-IDF</u>: The NOTAM text was analyzed using the TF-IDF (Term Frequency-Inverse Document Frequency) vectorization technique, resulting in feature vectors indicating the relevance of each term in the corpus.

<u>Justification</u>: TF-IDF weighting takes into account both a term's frequency within a text and its rarity across all documents, highlighting discriminative terms for each NOTAM.

<u>DistilBERT</u>: The DistilBERT model, a version of BERT (Bidirectional Encoder Representations from Transformers), was used to extract sentence embeddings from the NOTAM text.

<u>Justification</u>: DistilBERT generates contextualized embeddings for each sentence in the NOTAM text, encapsulating the semantic meaning of the complete sentence and allowing classifiers to use contextual information for classification.

<u>MPNet:</u> Sentence Embedding (MPNet) The MPNet (Miniature Pre-trained Transformer) model was used to generate sentence embeddings from the NOTAM text.

<u>Justification:</u> MPNet, like other transformer-based models, extracts contextual information from the full text sequence, resulting in rich embeddings that capture the semantic meaning and relationships of the NOTAM text.

Chapter 4 : Analysis

4.1 Text Handling Methods

4.1.1 TF-IDF

Term Frequency - Inverse Document Frequency (TF-IDF) is a popular statistical method for natural language processing and information retrieval. It assesses the importance of a phrase within a document in comparison to a collection of documents (i.e., a corpus).

A text vectorization procedure transforms words in a text document into numbers. There are number of text vectorization scoring techniques, with TF-IDF being one of the most used.

TF-IDF transformation was applied to the details_expanded column in the NOTAM dataset using Dataiku. The following customization settings were used:

- **Min. Rows Fraction** : 0.1 Words that don't appear in this fraction of rows (0.1%) will not be considered, helping filter out rare terms that may not contribute significantly to the analysis.
- **Max. Rows Fraction** : 80% Words that appear in more than this fraction of rows will not be considered, as they are likely to be too common and may not provide valuable information.
- **Stop Words**: No custom stop words were specified, allowing the default stop word list to be used.

4.1.2 Sentence Embedding using DistilBert

Sentence embedding with the DistilBERT pretrained model from Hugging Face entails converting textual sentences into fixed-size vector representations using the pre-trained DistilBERT architecture. DistilBERT is a lightweight BERT model that was trained on huge text datasets to learn contextualized representations of words and phrases. Using this pretrained model, textual phrases can be turned into dense vector embeddings, capturing their semantic meaning and helping downstream natural language processing tasks like semantic similarity calculation, text classification, and clustering. The "distilbert-base-uncased" variation was chosen for reasons such as model performance, computational efficiency, and compatibility with the NOTAM dataset.

4.1.3 Sentence Embedding using Sentence-Transformer-MPNet-v2:

This project's pretrained model for sentence embedding is the Sentence-Transformer-MPNet-v2, which is available on Hugging Face. This model uses the MPNet-v2 architecture and was pre-trained on huge text corpora to learn contextualized sentence representations. Using this pretrained model, textual phrases can be translated into dense vector embeddings, allowing for downstream tasks like semantic similarity calculation and text classification.

4.2 Data Preprocessing

4.2.1 Model Comparison

Different text processing techniques, including as TF-IDF, DistilBert, and mpnet, were used to train and assess a variety of machine learning models. The selected models are outlined below, along with each one's corresponding accuracy:

4.2.1.1 Random Forest (TF-IDF)

Random Forest was trained to identify the target variable from the TF-IDF transformed data. This technique uses the strength of numerous decision trees to give strong performance across a variety of datasets. The Random Forest method discovered the underlying patterns and relationships inside the TF-IDF features, reaching an accuracy of 0.74 (± 0.01) on the testing set. The model's precision, recall, and F1 score were all commendable, demonstrating its ability to categorize the target variable. Below are the settings used.

Setting	Min	Max	Description
Number of Trees	80	200	Number of trees in the forest.
Maximum Depth of Tree	6	20	Maximum depth of each tree. Higher values improve prediction but may cause overfitting and increase training time.
Minimum Samples per Leaf	1	20	Minimum samples required in a leaf node to split it. Lower values increase granularity but may lead to overfitting. Higher values promote generalization but may miss details.

Table 4.1: Random forest settings.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 70% of the NOTAMs that should be removed were accurately classified as such,

whereas 23% were shown to be kept. In contrast, 77% of the NOTAMs that should be kept were correctly classified, while 30% were predicted to be removed.

	Predicted remove	Predicted keep
Actually remove	70 %	23 %
Actually keep	30 %	77 %
Total	100 %	100 %

Table 4.2: Confusion matrix for Random Forest (TF-IDF).

Figure 4.1 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 81.2%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

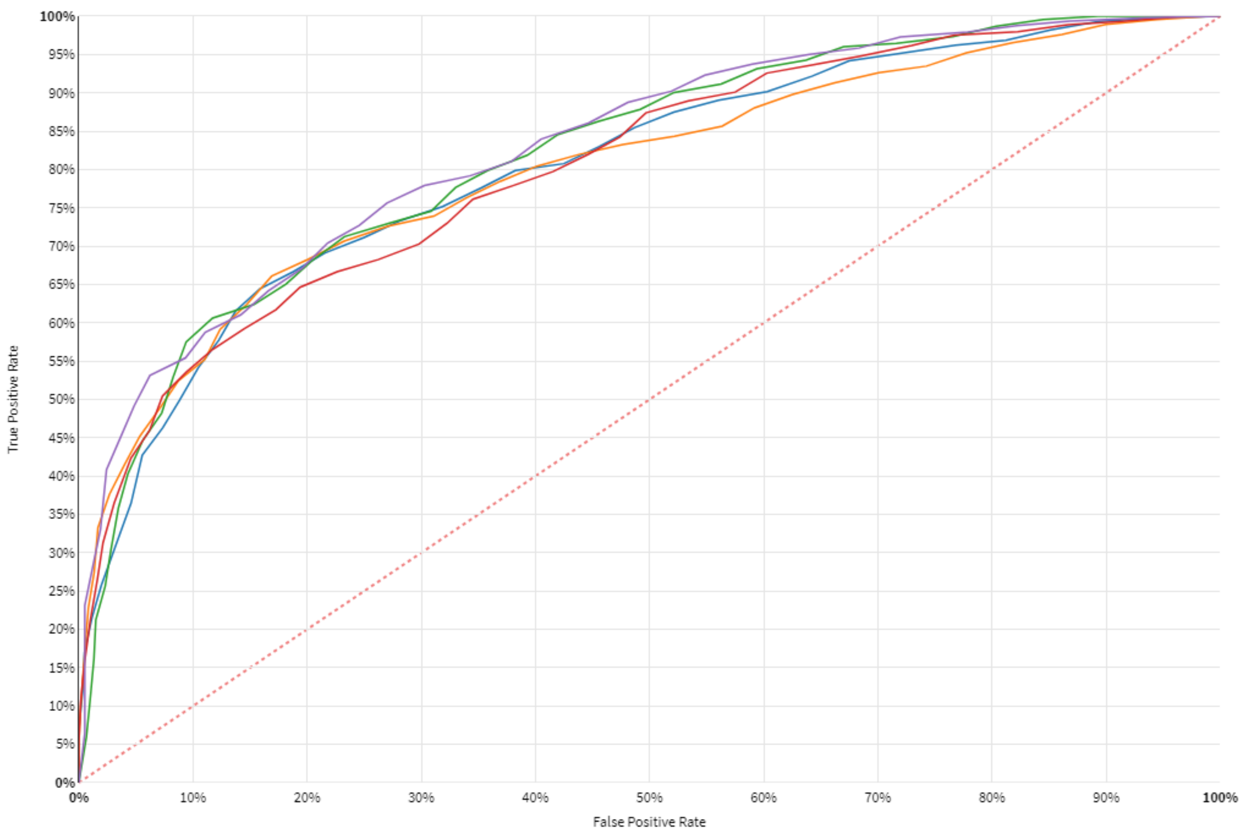


Figure 4.1: ROC curve for Random Forest *(TF-IDF)*

4.2.2.1 Random Forest (DistilBert)

With the same previous random forest setting used It was trained to categorize the target variable using DistilBert embeddings. DistilBert, a lightweight variation of the BERT technology, provides efficient performance without sacrificing precision. The Random Forest model trained on DistilBert embeddings has an accuracy of 0.76 (± 0.02) on the testing set, indicating its ability to extract contextual information from text input.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 69% of the NOTAMs that should be removed were accurately classified as such, whereas21% were shown to be kept. In contrast, 79% of the NOTAMs that should be kept were correctly classified, while 31% were predicted to be removed.

	Predicted `remove`	Predicted `keep`
Actually `remove`	69 %	21 %
Actually `keep`	31 %	79 %
Total	100 %	100 %

Table 4.3: Confusion matrix for Random Forest (DistilBERT).

Figure 4.2 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 83.1%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

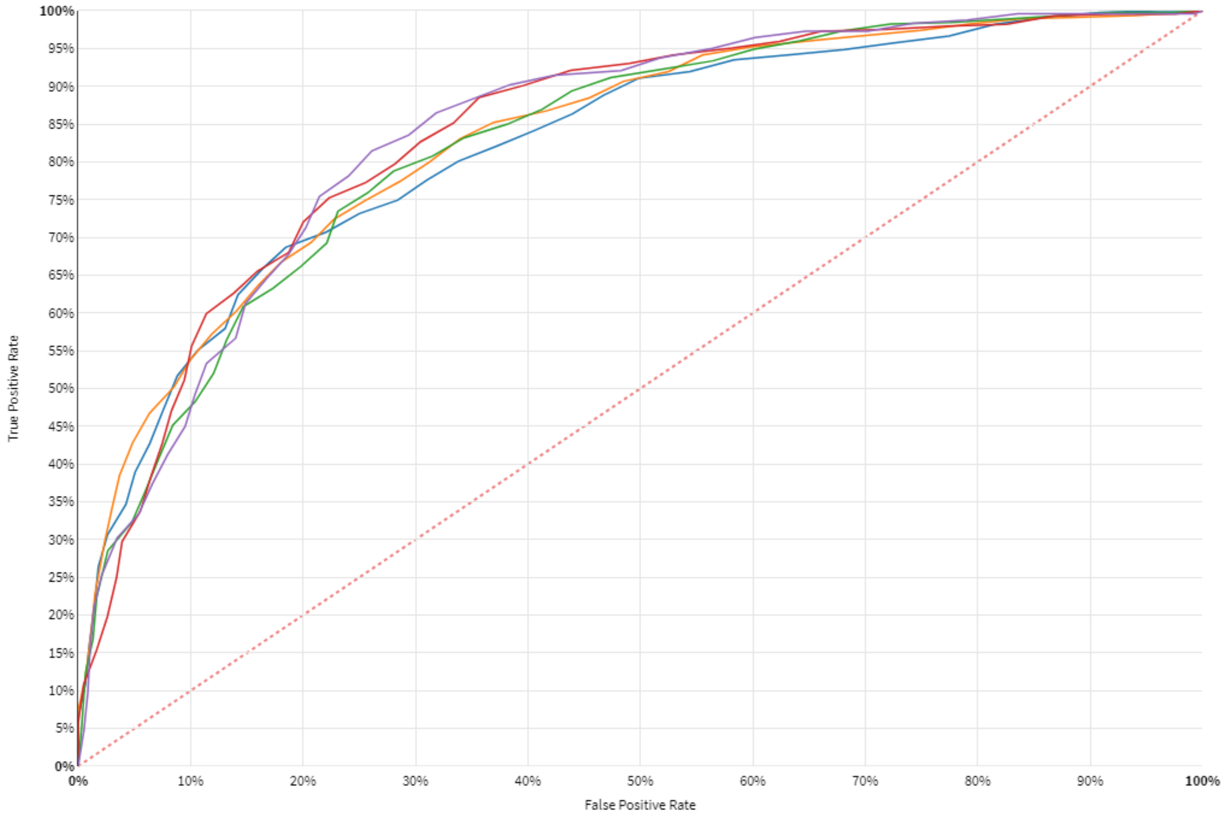


Figure 4.2: ROC curve for Random Forest (DistilBERT)

4.2.2.1 Random Forest (MPNet)

With the same previous random forest setting used It was trained to categorize the target variable using mpnet embeddings. MPNet, another transformer-based model, was used to create embeddings for the NOTAM text. MPNet used multi-task learning to improve textual data representations, allowing the Random Forest classifier to perform better.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 68% of the NOTAMs that should be removed were accurately classified as such,

whereas18% were shown to be kept. In contrast, 82% of the NOTAMs that should be kept were correctly classified, while 32% were predicted to be removed.

	Predicted `remove`	Predicted `keep`
Actually `remove`	68 %	18 %
Actually `keep`	32 %	82 %
Total	100 %	100 %

Figure 4.4: Confusion matrix for Random Forest (MPNet).

Figure 4.3 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 83.9%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

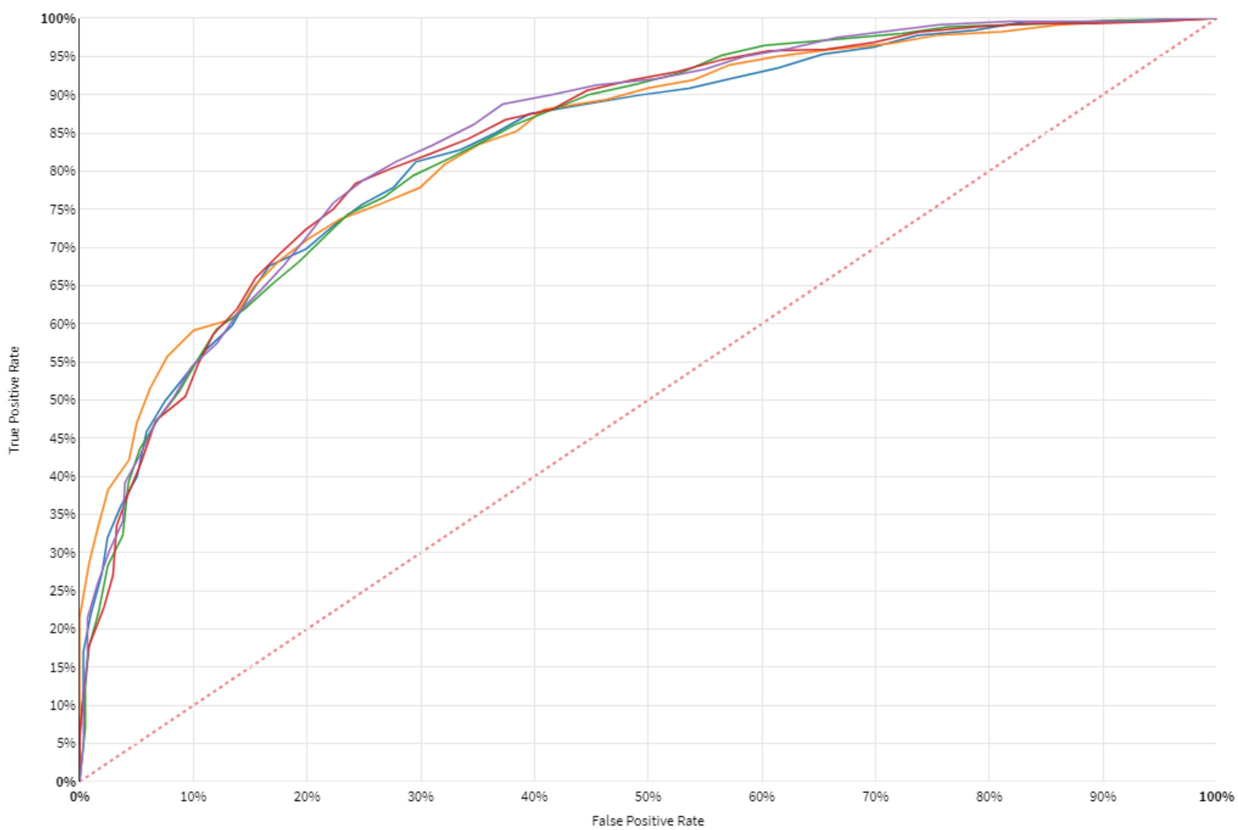


Figure 4.3: ROC curve for Random Forest (MP Net)

4.2.2.1 **Support Vector Machine** (TF-IDF)

Support Vector Machine (SVM) classifier was trained to classify NOTAMs using TF-IDF transformed data. The model performed well, with an overall classification accuracy of 76%.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 73% of the NOTAMs that should be removed were accurately classified as such, whereas 20% were shown to be kept. In contrast, 80% of the NOTAMs that should be kept were correctly classified, while 27% were predicted to be removed.

	Predicted `remove`	Predicted `keep`
Actually `remove`	73 %	20 %
Actually `keep`	27 %	80 %
Total	100 %	100 %

Table 4.5: Confusion matrix for SVM (TF-IDF).

Figure 4.4 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 84%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

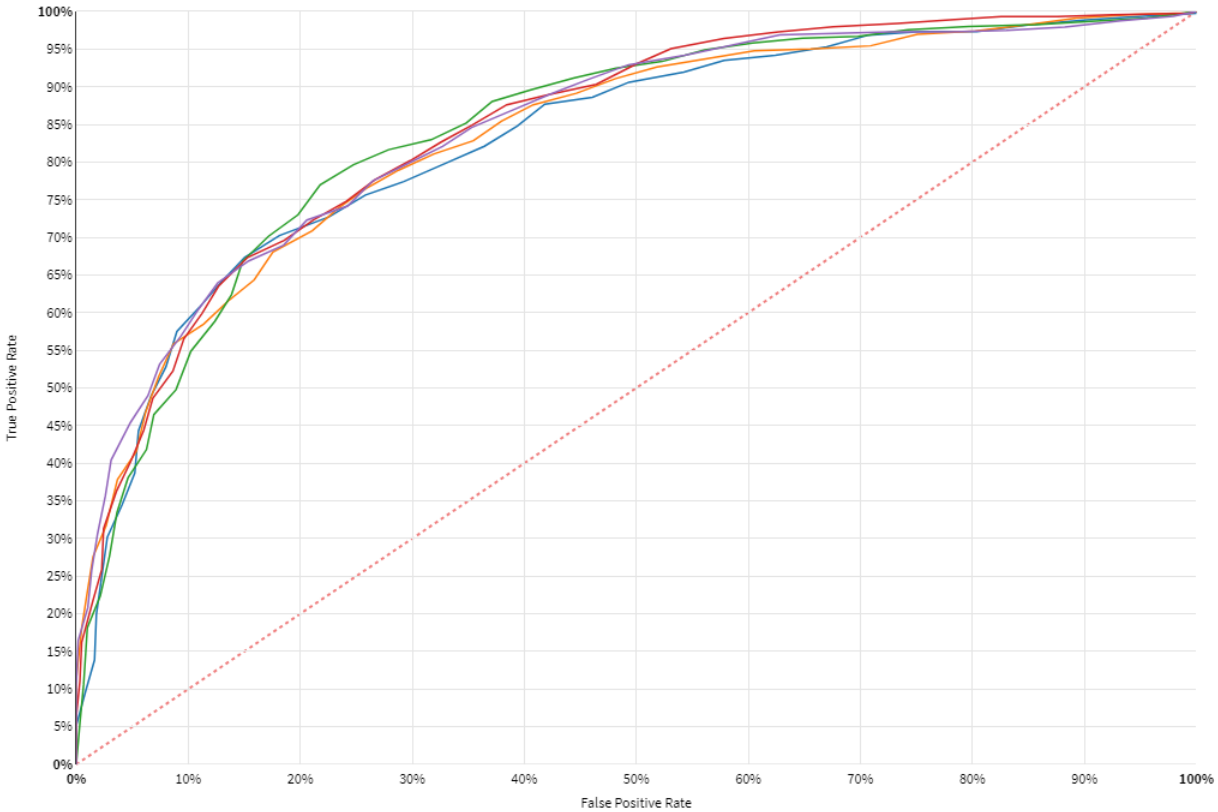


Figure 4.4: ROC curve for SVM (TF-IDF)

4.2.2.1 Support Vector Machine (DistilBert)

With the same previous SVM settings used It was trained to classify the target variable using DistilBert embeddings.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 75% of the NOTAMs that should be removed were accurately classified as such,

whereas22% were shown to be kept. In contrast, 78% of the NOTAMs that should be kept were correctly classified, while 25% were predicted to be removed.

	Predicted remove	Predicted keep
Actually remove	75 %	22 %
Actually keep	25 %	78 %
Total	100 %	100 %

Table 4.6: Confusion matrix for SVM(DistilBERT)

Figure 4.5 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 82.4%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

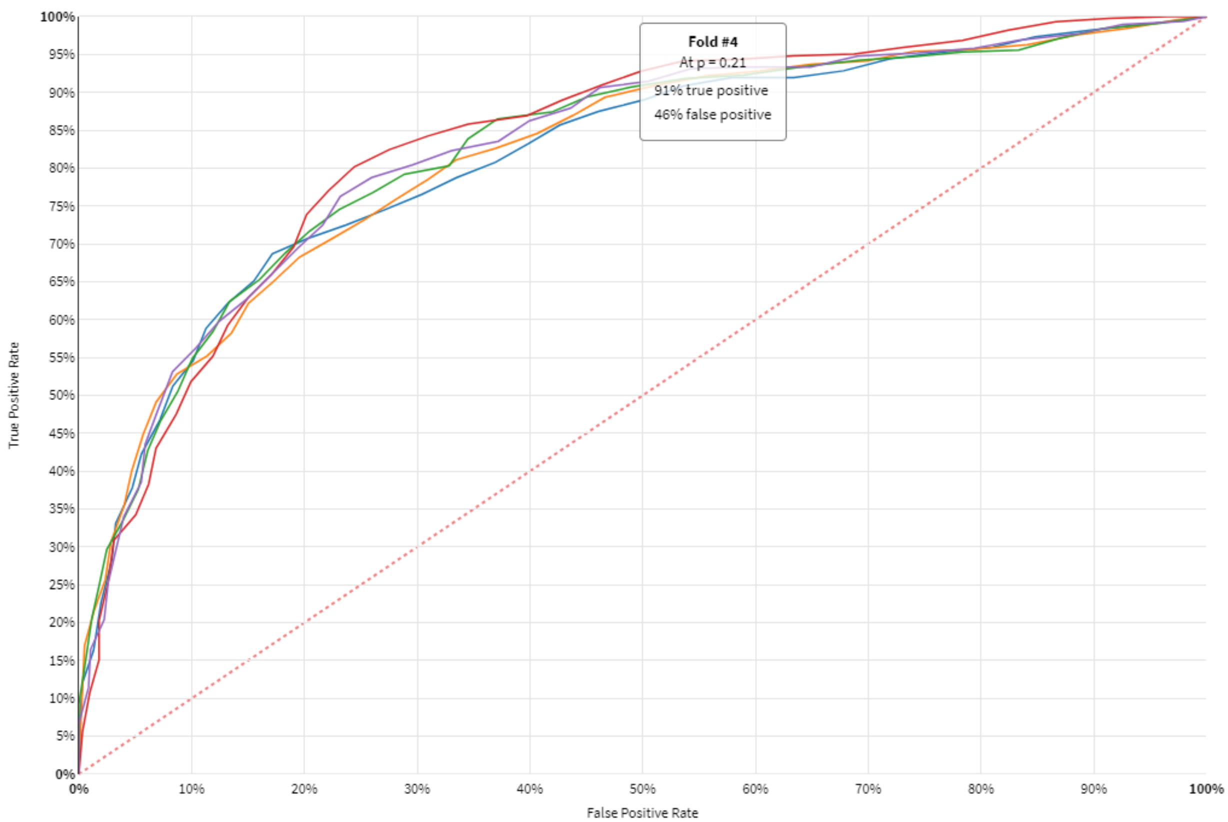


Figure 4.5: ROC curve for SVM (DistilBERT)

4.2.2.1 **Support Vector Machine** (MPNet)

The SVM model trained on mpnet embeddings attained an accuracy of 0.71 on the testing set, indicating its capacity to detect subtle patterns within text data.

The confusion matrix shows the model's classification performance for NOTAM data. It indicates that 68% of the NOTAMs that should be removed were accurately classified as such, whereas19% were shown to be kept. In contrast, 81% of the NOTAMs that should be kept were correctly classified, while 32% were predicted to be removed.

	Predicted `remove`	Predicted `keep`
Actually `remove`	68 %	19 %
Actually `keep`	32 %	81 %
Total	100 %	100 %

Table 4.7: Confusion matrix for SVM (MPNet).

Figure 4.6 below visualizes the Receiver Operating Characteristic (ROC) curve. The Area under the Curve (AUC) is 83.9%, which is close to 1. This indicates that the model has a very good ability to discriminate between relevant NOTAMs and irrelevant ones.

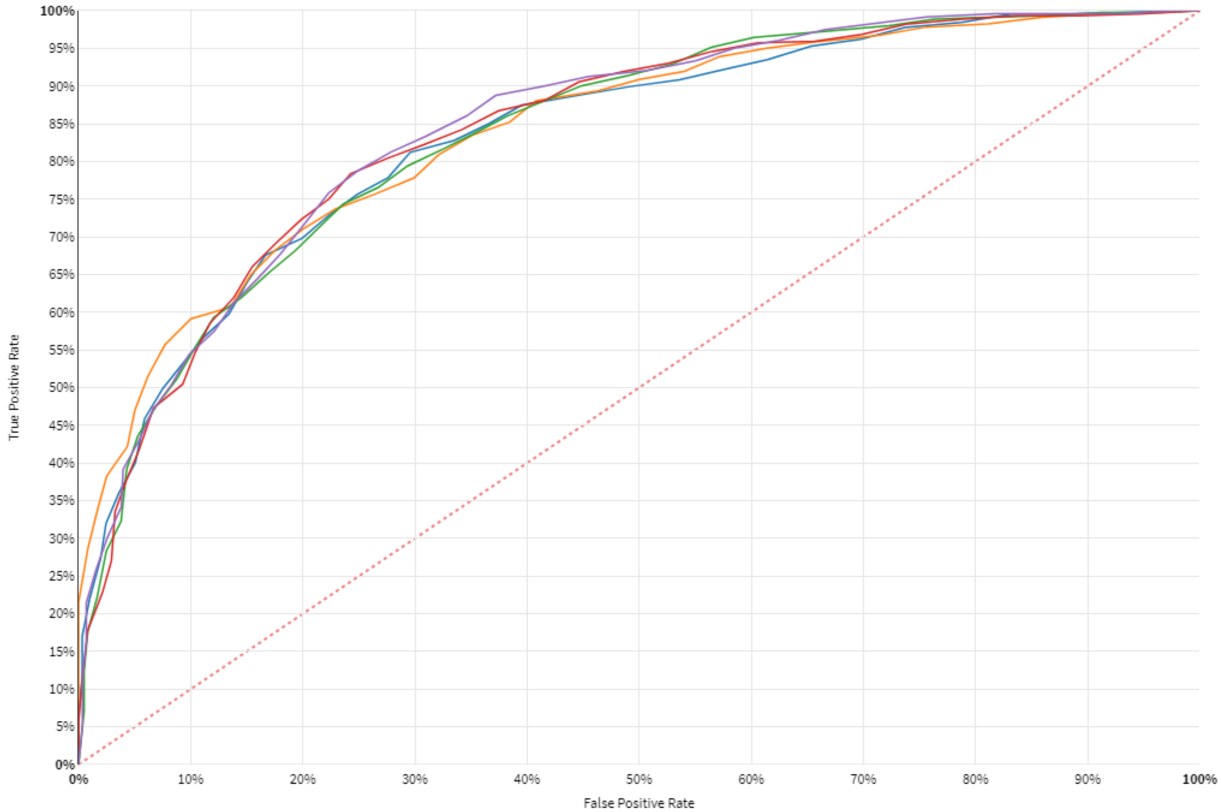


Figure 4.6: ROC curve for SVM (MPNet)

Model	Train Time	Accuracy	Precision	Recall	F1 Score	ROC
Random Forest (TF-IDF)	25s	0.74	0.76	0.71	0.75	0.81 (± 0.02)
SVM (TF-IDF)	23m 45s	0.76	0.72	0.73	0.73	0.84 (± 0.01)
Random Forest (DistilBert)	4m 15s	0.75	0.71	0.73	0.72	0.82 (± 0.01)
SVM (DistilBert)	8m 25s	0.76	0.74	0.71	0.72	0.82 (± 0.01)
Random Forest (MPnet)	7m 52ss	0.71	0.78	0.68	0.73	0.84 (± 0.01)
SVM (MPNet)	11m 24ss	0.72	0.82	0.70	0.75	0.84 (± 0.02)

Table 4.8: comparing performance metrics.

In summary, using the Support Vector Machine with TF-IDF produced the best results across a variety of metrics, including accuracy and ROC.

Chapter 5: Conclusions and Recommandations

In conclusion, this chapter summarizes the study's important findings, draws insightful inferences, and makes recommendations for future research endeavours. The major goal of this research was to create a machine learning-based model that uses NOTAM text to properly predict the decision whether to keep or remove NOTAM from an OFP, solving persisting issues in undergoing manual review on all OFPs. Using the provided NOTAM dataset from an aviation organization, the study methodically classified a large number of NOTAM entries, allowing for extensive analysis.

The preprocessing phase included cleaning up the dataset and standardizing the format for consistency. It also included expanding aviation abbreviations.

For sampling, a class rebalance strategy with an estimated ratio of 100% for the column "decision" was used to ensure representation across classes. A fixed random seed (1337) was used to assure sampling reproducibility.

The proposed Support Vector Machine model using TF-IDF for text handling correctly classified 76% of NOTAMs. Comparative analysis against five other machine learning classifiers revealed the superiority of the proposed approach over other Machine models and text handling approaches.

This research adds substantial insights to the creation of an effective NOTAM prediction system, providing valuable assistance in improving aviation safety and operating efficiency. The results of rigorous experimentation demonstrate the suggested model's reliability in forecasting relevant NOTAM entries, which is critical for pilots, air traffic controllers, and aviation stakeholders. Furthermore, the proposed approach is feasible in real-world circumstances, demonstrating its ability to streamline NOTAM processing and decision-making procedures.

However, continual monitoring and refining of the model are required to suit changing aviation requirements and assure long-term accuracy in NOTAM prediction.

5.4 Recommendations

The study developed a strong and reliable mechanism for filtering relevant and irrelevant NOTAMs. This study's findings provide vital insights into improving the management and understanding of NOTAMs, addressing significant difficulties in the aviation sector. This study's use of advanced machine learning algorithms has yielded encouraging results in properly forecasting NOTAM-related decisions contributing to better aviation safety and efficiency. Nonetheless, there is a research gap that highlights the need for future studies to investigate and develop techniques to improving NOTAM prediction. Further study is necessary to develop more advanced and trustworthy approaches adapted to the special needs and complexities of NOTAM data.

In this regard, Mechanisms for dynamic model updating and adaptation must be put in place in order to respond to the changing and dynamic nature of NOTAM entries and regulatory changes in the aviation sector. To maintain the highest possible accuracy over time, this involves automatically recalibrating features and retraining the predictive model with new data on a regular basis. The model can successfully capture changes in patterns and trends within NOTAM data by being updated with the most recent information. This keeps the model relevant and accurate in anticipating aviation-related events and risk.

Secondly, The seamless integration of predictive models into existing aviation systems is a vital step in maximizing their use for NOTAM prediction. The created predictive model can be integrated into established aviation systems such as flight planning software and air traffic management systems to provide pilots and air traffic controllers with real-time NOTAM predictions and decision support. This connection provides aviation professionals with fast access to predictive information, allowing them to make informed decisions and take proactive steps to avoid potential risks and interruptions. Finally, the seamless integration of predictive models into aviation systems improves operational efficiency and safety across a wide range of aviation operations, contributing to the industry's overall resilience and effectiveness.

Finally, Fine-tuning the classifier head could significantly improve the model's performance, especially with the Notam model. By fine-tuning the classifier head, the model may better react to the intricacies and specifics of Notam data, increasing its accuracy and efficacy in classifying Notam-related material.

References

[1] Butler, M. (2022, December 29). What Are NOTAMs in Aviation? Fluix. https://fluix.io/blog/notams-aviation

[2] Rosa. (2023, June 12). What is a NOTAM and what is it for? - Grupo One Air. Grupo One Air. https://www.grupooneair.com/what-is-a-notam/

[3] Federal Aviation Administration. (2020). NOTAMs: Back to Basics: Pilots. https://www.faasafety.gov/files/notices/2019/Feb/Pilots_NOTAM_primer_1-28-19.pdf

[4] Lufthansa Systems. (2021, October 5). NOTAM processing for the Lido Flight Planning Solutions (Part 1 of 2). https://www.lhsystems.com/blog-entry/notam-processing-lido-flight-planning-solutions-part-1-2

[5] Szeto1, A., & Das, A. (n.d.). *Classification of Notices to Airmen using Natural Language Processing*. NASA. https://ntrs.nasa.gov/api/citations/20230017551/downloads/20230017551_SciTech_2024_FINAL.pdf

[6] Coupon, J. (n.d.). Classification of NOTAMs for SWISS International Airlines us. Constructor.org. Retrieved December 24, 2023, from https://academy.constructor.org/final-projects/swiss

[7] M. Banane and A. Erraissi, "A comprehensive study of Natural Language processing techniques Based on Big Data," 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 2022, pp. 1492-1497, doi: 10.1109/DASA54658.2022.9765270.

[8] Amin, N., Yother, T. L., Johnson, M. E., & Rayz, J. T. (2022). Exploration of Natural Language Processing (NLP) applications in aviation. *Collegiate Aviation Review*, *40*(1). https://doi.org/10.22488/okstate.22.100211

[9] Mirończuk, A., & Protasiewicz, J. (2018). A recent overview of the state-of-the-art elements of text classification. Expert Systems With Applications, 106, 36–54.

[10] Kobus, C. (2022). Knowledge extraction from aeronautical messages (NOTAMs) with self-supervised language models for aircraft. ResearchGate. https://www.researchgate.net/publication/362024577_Knowledge_extraction_from_aeronautical_messages_NOTAMs_with_self-supervised_language_models_for_aircraft_pilots

[11] Yang, C., & Huang, C. (2023). Natural Language Processing (NLP) in Aviation Safety: Systematic Review of Research and Outlook into the Future. Aerospace, 10, 600. https://doi.org/10.3390/aerospace10070600

[12] Clarke, S. S. B., Maynard, P., Almache, J. A., Kumar, S. G., Rajkumar, S., Kemp, A. C., & Pai, R. (2023). Natural Language Processing Analysis of Notices To Airmen for Air Traffic Management Optimization. National Aeronautics and Space Administration.

[13] Burgstaller, F., Steiner, D., Schrefl, M., Gringinger, E., Wilson, S., & van der Stricht, S. (2015). AIRM-based, fine-grained semantic filtering of notices to airmen. 2015 Integrated Communication, Navigation and Surveillance Conference (ICNS), Herdon, VA, USA, pp. D3-1-D3-13. doi: 10.1109/ICNSURV.2015.7121222

[14] Prokhorov, A. V. (2022). Impact of NOTAM on security and efficiency performance of flights (overview). Civil Aviation High TECHNOLOGIES, 25(1), 21–34. https://doi.org/10.26467/2079-0619-2022-25-1-21-34

[15] Tabassum1, A., & Patil, R. (2020). A Survey on Text Pre-Processing & Feature Extraction Techniques in Natural Language Processing. *International Research Journal of Engineering and Technology (IRJET)*, *07*(06).

[16] Qisong, H., Zixuan, W., Wen, Y., Minrui, C., Guoqiang, C., & Yang, Y. (2023). Research on NOTAM Information Extraction of Civil Aviation with NLP. 2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Dali, China, pp. 520-523. doi: 10.1109/ICCASIT58768.2023.10351768

[17] Khan, S. (2021, December 11). *BERT, RoBERTa, DistilBERT, XLNet — which one to use?* Medium. https://towardsdatascience.com/bert-roberta-distilbert-xlnet-which-one-to-use-3d5ab82ba5f8

[18] S. Behera, N. Kolipakula, P. Swetha, Y. Yogish and J. R. Prathuri, "Text Classification using Machine Learning Algorithms on Encrypted Data," 2023 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/CONECCT57959.2023.10234826.

[19] Winastwan, R. (2024, April 6). *All-Mpnet-Base-V2: Enhancing Sentence Embedding with AI - Zilliz blog*. Zilliz. https://zilliz.com/learn/all-mpnet-base-v2-enhancing-sentence-embedding-with-ai#Understanding-MPNet

[20] Tabsharani. (n.d.). *What is a support vector machine (SVM)?* TechTarget. Retrieved February 2024, from https://www.techtarget.com/whatis/definition/support-vector-machine-SVM

[21] Z. Liu, X. Lv, K. Liu and S. Shi, "Study on SVM Compared with the other Text Classification Methods," 2010 Second International Workshop on Education Technology and Computer Science, Wuhan, China, 2010, pp. 219-222, doi: 10.1109/ETCS.2010.248.

[22] Ohata, E. F., Mattos, C. L. C., Gomes, S. L., De S Rebouças, E., & Rego, P. a. L. (2022). A text classification methodology to assist a large technical support system. *IEEE Access, 10,* 108413–108421. https://doi.org/10.1109/access.2022.3213033

[23] S. Karamizadeh, S. M. Abdullah, M. Halimi, J. Shayan and M. j. Rajabi, "Advantage and drawback of support vector machine functionality," 2014 International Conference on Computer, Communications, and Control Technology (I4CT), Langkawi, Malaysia, 2014, pp. 63-65, doi: 10.1109/I4CT.2014.6914146.

[24] *What is Random Forest? | IBM*. (n.d.). https://www.ibm.com/topics/random-forest

[25] Tin Kam Ho, "Random decision forests," Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 1995, pp. 278-282 vol.1, doi: 10.1109/ICDAR.1995.598994.

[26] International Civil Aviation Organization (ICAO). (2010). ICAO Abbreviations and Codes: Procedures for Air Navigation Services, Eighth Edition.

[27] Hugging Face, "distilbert-base-uncased" Hugging Face, [Online]. Available: https://huggingface.co/distilbert/distilbert-base-uncased. [Accessed 2024].

[28] Hugging Face, "all-mpnet-base-v2" Hugging Face, [Online]. Available: https://huggingface.co/sentence-transformers/all-mpnet-base-v2 [Accessed 2024].

[29] Hoque, M. (2023, May 1). *A Comprehensive Overview of Transformer-Based Models: Encoders, Decoders, and More*. Medium. https://medium.com/@minh.hoque/a-comprehensive-overview-of-transformer-based-models-encoders-decoders-and-more-e9bc0644a4e5

[30] Hamdaoui, Y. (2024, February 8). *A Guide to TF-IDF*. Built In. https://builtin.com/articles/tf-idf

[31] R, S. E. (2024, April 19). *Understand Random Forest Algorithms With Examples (Updated 2024)*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/

[32] OMAE Valid NOTAMs as of 07 MAY 2024 2120 UTC. (2024, May 7). Retrieved from https://www.gcaa.gov.ae/en/ais/Pages/notam.aspx#