Enhancing Ambulance Resource Management through Machine Learning-based Demand Prediction in Dublin, Ireland

MSc Research Project
MSc Data Analytics

Ciarán Jones
Student ID: 21135231

School of Computing
National College of Ireland

Supervisor: Anh Duong Trinh

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	……Ciarán Jones……………………………………………………………………………
Student ID:	……21135231…………………………………………………………………………….
Programme:	…… MSc in Data Analytics ……………………… **Year:** ……2023…………..
Module:	……… MSc Research Project………………………………………………….………
Supervisor:	……… Anh Duong Trinh ………………………………………………………….………
Submission Due Date:	………14/08/2023………………………………………………………………….………
Project Title:	… Enhancing Ambulance Resource Management through Machine Learning-based Demand Prediction in Dublin, Ireland……….………
Word Count:	……8236………………… **Page Count**…...25……………………………

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……………………………………………………………………………………………………

Date: ………14/08/2023…………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Ambulance Resource Management through Machine Learning-based Demand Prediction in Dublin, Ireland

Ciarán Jones
21135231

Abstract

This study investigates the efficacy of machine learning models – XGBoost, Random Forest (TensorFlow Decision Forest), and MLP Neural Network – in predicting ambulance demand in Dublin city. This study, aimed at improving resource allocation within emergency medical services (EMS), bridges gaps in the current understanding of Dublin's ambulance demand dynamics. Utilizing a comprehensive dataset comprising over 850,000 historical ambulance demand records over ten years, we scrutinized the influence of diverse feature engineering techniques on the models' performance. The study primarily accomplishes three key contributions: (1) Unveiling an effective XGBoost model, coupled with astute feature selection, for ambulance demand prediction, which achieved a mean absolute error (MAE) of 0.49673, thereby contributing to strategic EMS planning; (2) Highlighting significant factors that influence ambulance demand, notably temporal and societal factors, while contesting the previously assumed importance of weather data; and (3) Underscoring the essential role of feature engineering in refining model performance. Our findings suggest potential areas of improvement in model performance, through further refinement and integration of additional data sources. This paves the way for future research to enhance these models and assess their applicability across different regions, ultimately augmenting EMS resource allocation and public health outcomes.

Keywords: ambulance demand, machine learning, predictive modelling, Dublin, resource management, public health, feature engineering.

1 Introduction

Emergency medical services (EMS) are pivotal in delivering immediate medical care to communities, with a primary objective of reducing response times and the rate of death and morbidity (Aringhieri et al., 2017). EMS managers and dispatchers need to understand the distribution of incoming call requests (demand) and develop resource deployment plans based on historical demand data and forecasts. This necessitates accurate emergency call forecasts, also known as ambulance demand prediction.

The crucial significance of Emergency Medical Services (EMS) is emphasized through its function in enhancing survival chances by administering immediate care to those experiencing critical emergencies (Jin et al., 2021). In contrast, the unequal distribution of EMS supply and demand in urban regions might lead to a scarcity of readily available EMS resources, consequently postponing the initial medical assistance. This situation underscores the immediate necessity to unearth the concealed relationship between EMS supply and demand, forecast upcoming EMS requirements, and implement measures to prepare for unforeseen emergencies.

In Dublin, Ireland, the Health and Social Care Executive (HSE) ambulance service is experiencing a surge in ambulance demand, straining the system (National Ambulance Service, 2020). Factors such as an aging and expanding population requiring more medical care are challenging the government's resource management capabilities. In 2020, the pressure on emergency services led to prolonged delays and reduced satisfaction among patients. Specifically, over 40% of urgent emergency calls surpassed the targeted 20-minute response window (National Ambulance Service, 2020).

To tackle this problem, the current research is focused on crafting a machine learning (ML) model that can precisely forecast the need for ambulances in Dublin. Building upon prior studies that employed ML techniques for predicting ambulance requirements, this research seeks to identify the most suitable ML approach for the Dublin area. Additionally, the investigation includes an examination of how newly considered factors, such as public holidays, influence the efficiency of the ML model.

The exploration fills a notable research void concerning ambulance demand forecasting in Ireland, and evaluates models to determine the optimal one. By fashioning an accurate and trustworthy ML framework to anticipate ambulance needs, the National Ambulance Service (NAS) stands to refine its handling of resources and assure the prompt and suitable delivery of medical aid. Moreover, this research contributes to the wider academic domain by analysing the effects of supplemental factors on the prediction of ambulance demand. The anticipated findings of this study hold substantial promise in augmenting the comprehension and prediction of ambulance requirements in Ireland, paving the way for enhanced medical services and improved patient health outcomes.

The research questions guiding this study are:

What are the optimal machine learning methods for predicting ambulance demand in Dublin, Ireland, and how do they compare in terms of performance?

To what extent can the performance of ambulance demand prediction models be improved with feature engineering and the addition of new features, and how do these improvements vary across different prediction models?

The paper is structured as follows: **Section 2** reviews existing literature on ambulance demand prediction, including machine learning models and feature engineering. **Section 3** details this studies methodology—data collection, pre-processing, model development, and evaluation metrics. **Section 4** presents the design specification of the prediction system. **Section 5** outlines the implementation, detailing data transformation, tools used, and models implemented. **Section 6** evaluates and compares the performance of XGBoost, Random Forest, and MLP Neural Network models, also discussing the role of feature engineering. Finally, **Section 7** concludes the study, summarizing the findings and suggesting potential future work.

2 Related Work

2.1 Ambulance Demand in Dublin, Ireland

In Dublin, Ireland, the growing need for pre-hospital emergency services has been triggered by an increasingly aging population and a boost in chronic illness (NAS, 2020). The Irish National Ambulance Service (NAS), designed as a response-driven service, strives to serve the entire populace. Meeting the needs of a constantly changing population presents continuous challenges for NAS, especially considering the demographic trend towards aging. The NAS works in coordination with the Sláintecare Programme Implementation Office within the Health Department, aligning their strategies with the National Service Plan (NSP) to ensure effective care delivery in the mid-term (NAS, 2020). Predictive techniques could facilitate the optimal distribution of resources and cut down wait times, ensuring immediate provision of health and social care according to clinical necessities (Kyrkou et al., 2023). Nevertheless, the current academic discourse lacks substantial insights into the utilization of predictive models specifically aimed at ambulance demand in Dublin, Ireland.

2.2 Predicting Ambulance Demand

Traditional prediction methods of ambulance requirements have relied on classic time series methodologies, such as autoregressive and ARIMA models, as corroborated by the work of Baker & Fitzpatrick (1986) since the 1980s. Despite their prevalence, these models have inherent drawbacks, chiefly their inability to navigate complex, non-linear correlations often found in data relating to ambulance demand, a fact noted by Hyndman & Athanasopoulos (2018). Recently, a trend has emerged to harness machine learning technologies to overcome these limitations. For example, Ramgopal et al. (2021) found that employing machine learning techniques like XGBoost and decision trees could enhance the precision of ambulance demand predictions. This has led the academic field to lean towards machine learning methodologies to develop more reliable and efficient models for forecasting ambulance needs.

Several comparative research papers have evaluated various predictive models for ambulance demand, generally highlighting the superiority of machine learning models over conventional time series methods. The pioneering comparative research by Setzler et al. (2009) pitted machine learning models, such as ANN, against traditional models like ARIMA and seasonal naïve methods, demonstrating the higher accuracy of machine learning models, particularly ANN.

In a similar vein, Martin et al. (2021) assessed the effectiveness of machine learning models, including artificial neural networks, against standard time series techniques like ARIMA, Holt-Winters, and MHF, showing that ANN models excel in capturing intricate connections between predictive factors and ambulance requirements.

Furthermore, Ramgopal et al. (2021) conducted an analysis comparing various machine learning models like XGBoost, LSTM, and MLP for predicting ambulance demand, concluding that XGBoost and Random Forest outperformed others in accuracy and efficiency.

The literature consistently indicates that machine learning models, especially XGBoost, random forest, and artificial neural networks, are more adept at forecasting ambulance needs in comparison to traditional approaches. However, the choice of the best machine learning model may differ based on the specific data and features used in the model, making it crucial to carefully analyse and weigh multiple models to select the one that best aligns with the requirements (Wickramasuriya et al., 2019).

2.2.1 Important Features in Predicting Ambulance Demand

The forecast of ambulance needs is an intricate procedure that entails multiple essential features. These can be grouped into three main categories: time-related, location-related, and external aspects.

Time-related factors reflect the changing ambulance needs over various time frames. Numerous studies have identified significant time-based elements such as the hour of the day, specific days within the week, particular months, and seasonal variations. For example, Ramgopal et al. (2019) noticed an upsurge in emergency calls on weekends and holidays as opposed to regular weekdays and an increased demand in winter compared to summer. Still, certain variables like public holidays and substantial communal events that may influence ambulance needs remain largely unexamined. Moreover, Lin et al. (2020) brought forward the idea of accumulated counts or aggregate demand over the preceding 7 and 30 days as a possible time-related component, providing a historical framework that could boost the predictive model's precision.

Location-related aspects are crucial since they explain the varying demands across different geographical areas. Elements such as the site of the incident, population concentration, and proximity to the closest medical facility play an important role in forecasting ambulance needs (Ramgopal et al., 2019). Previous research has detected a robust connection between population concentration and the density of calls, as emphasized by Vile et al. (2016) and Chen et al. (2016). Hence, the inclusion of location-related aspects in the model for predicting ambulance demand is vital for achieving accuracy.

External aspects, although not directly connected to ambulance needs, can affect it. Factors like climatic conditions, extraordinary events, and demographic shifts have been identified as influential. Research has shown that weather variations such as temperature, precipitation, and snow can considerably affect ambulance needs (Ramgopal et al., 2019). The integration of external aspects like temperature, wind velocity, and atmospheric pressure has been found to enhance predictive accuracy by McLay, Boone, and Brooks (2012). Nevertheless, studies focusing on EMS needs have largely omitted aspects like public holidays, leaving room for more inclusive research encompassing a broader set of features.

In summary, the forecast of ambulance needs requires attention to various crucial components, categorized as time-related, location-related, and external aspects. While significant time and location factors have been investigated, there exists a demand for more thorough research, embracing an extended selection of components, such as public holidays, to improve the precision of ambulance demand predictions. The exploration of novel time-related aspects like accumulative counts highlights promising avenues for future studies in this domain.

2.3 Research Niche and Expected Contribution

Despite advancements in machine learning applied to predicting ambulance needs, there remains a deficiency in studies focusing on its implementation, particularly in Dublin, Ireland. Hence, this research is set to explore the utilization of machine learning techniques in forecasting ambulance demand in this region, with an emphasis on pinpointing the most accurate algorithm and comprehending the factors that shape ambulance demand.

This investigation aims to enrich the existing scholarly corpus by incorporating elements like climatic conditions and public holiday events into the model predicting ambulance demand. It also plans to examine the potential of newly identified time-related features such as accumulative counts. The goal is to evaluate whether these variables have a substantial influence on ambulance needs and if they should be considered in the development of predictive models.

The anticipated contribution of this inquiry is a refined and more inclusive model for forecasting ambulance demand, enabling emergency health services to distribute resources more proficiently and enhance responsiveness. Moreover, the results of this study could be applied to other urban settings sharing resemblances with Dublin, Ireland.

Ultimately, this study endeavours to offer a substantial contribution to the realm of ambulance demand forecasting, specifically in Dublin, Ireland, and more generally in the wider field of prediction. This will be achieved by addressing existing research gaps and presenting a more holistic and accurate predictive model.

3 Research Methodology

The Cross-Industry Standard Process for Data Mining (CRISP-DM) was employed in this study due to its structured, iterative nature that facilitates continuous improvement (Larose & Larose, 2015). This versatile, industry-standard framework guides us through all stages of a data mining project, making it suitable for a wide variety of data types and business challenges. It ensures a comprehensive process of data exploration, preparation, modelling, and evaluation. Each phase of the CRISP-DM methodology will be further discussed in Section 4, providing an in-depth understanding of its application in this research.

3.1 Data Collection and Pre-processing

To develop the machine learning models for predicting ambulance demand, historical ambulance call-out data was obtained from the National Ambulance Service (NAS) in Dublin for a 10-year period from 2013 to 2022. This public dataset was obtained via the Smart Dublin Website (Smart Dublin, 2023). This dataset includes time-stamped records of each ambulance call-out, along with the location of the call-out (represented as Station ID), call priority (e.g., life-threatening, non-life-threatening), and dispatch and arrival times.

Table 1: Overview of the Ambulance Call-out Dataset

Feature	Description	Count	Unique	Most Frequent	Frequency
Date	When the incident was logged	846630	70304	01/01/2014	304
Station Area	DFB Station Boundary where the incident occurred	846630	17	Tallaght	137031
Clinical Status	Acuity of the reported incident (from Echo to Omega)	846630	8	Delta	375022
TOC	Time of Call	846630	230086	19:17:15	21
ORD	Time first appliance is MOBILISED or ORDERED to incident	846630	153483	17:53:36	29
MOB	Time the first appliance is MOBILE to INCIDENT	828438	151926	17:59:09	26
IA	In Attendance - time the first appliance is at the scene of the incident	704393	143180	17:54:33	25
LS	Leaving Scene (ambulance calls only)	527104	132859	10:43:54	20
AH	At Hospital (ambulance calls only)	556396	133887	20:16:49	21
MAV	Mobile and Available	772885	149533	17:35:40	31

Apart from the ambulance dispatch data, several other pertinent datasets were gathered, comprising:

- Meteorological data sourced from Met Éireann (Met Eireann, 2023), the Irish National Meteorological Service, incorporating temperature, precipitation, wind velocity, and humidity.
- Public holiday data procured from timeanddate.com (Time&Date.com, 2023), which was subsequently formatted into a dataset utilizing Python.

The pre-processing stage of the data entailed various steps, including:

- Data cleaning to eliminate any inaccuracies or inconsistencies.
- Handling missing values, such as imputation or exclusion based on the degree of missingness and their impact on the analysis.
- Identifying and dealing with outliers, which can significantly affect model performance and predictions.
- Organizing the primary dataframe by the station area.

- Merging the distinct datasets predicated on the date and time of each ambulance dispatch.
- Consolidating the data into hourly intervals to furnish a suitable level of detail for the analysis.
- Converting categorical variables into numerical form through one-hot encoding or other appropriate techniques.
- Formulating new features to incorporate cumulative counts of dispatches and other relevant temporal patterns.
- Scaling the numerical variables with the StandardScaler or other normalization techniques to guarantee they all fall within a comparable range.

3.2 Feature Selection

The process of feature selection plays an instrumental role in establishing a proficient predictive model. The features contemplated in this study were extracted from the data that had been collected and pre-processed as detailed earlier. These encompass temporal elements (hour of the day, day of the week, month, season), spatial elements (station area), and external elements (weather conditions, public events).

An analysis of feature importance was carried out to ascertain the relative significance of each feature in the prediction of ambulance demand. The feature importance was determined through the application of the XGBoost algorithm's feature importance methodology. This methodology utilizes a metric known as 'F-score' or 'Gain'. The XGBoost algorithm calculates the relative importance of a feature based on the number of times a feature is used to split the data across all trees, weighted by the improvement to the model as a result of each split. The feature that provides the most value for splitting the data, thus improving the model's accuracy, is considered the most important.

The more often a feature is used in the trees of the model, and the more it improves the model's performance when it is used, the higher its relative importance or F-score. This score provides an indication of the contribution each feature makes to the model's predictive power. Features with a higher F-score are considered more important for prediction.

This F-score, or gain, is not only an efficient way of selecting the most relevant features but also helps in reducing the dimensionality of the dataset, minimizing the risk of overfitting, and improving the model's overall performance.

Applying the XGBoost feature importance methodology in this study, we were able to prioritize features that significantly influence ambulance demand, ensuring a more robust and accurate prediction model. The specific results of this feature importance analysis will be elaborated upon in the following sections of the study.

3.3 Model Development

To predict ambulance demand, we employed a diverse set of machine learning algorithms, including Random Forest, Extreme Gradient Boosting (XGBoost), and Deep Learning models (MLP). These models were implemented using Python and relevant libraries, such as Scikit-learn, XGBoost, TensorFlow, and TensorFlow Decision Forests (TF-DF).

The model development process comprised the following key steps, optimized for reliable and accurate predictions:

1. **Data Splitting:**
   The dataset was divided into three distinct sets while preserving the chronological order essential for time-series data. The training set encompassed 70% of the total data, while the testing set comprised 20%. The validation set constituted the remaining 10%. This division ensured that the models were trained on historical data, validated on recent data, and tested on unseen data to gauge their generalization capabilities accurately.

2. **Hyperparameter Tuning:**
   For each model, we performed hyperparameter tuning to optimize their performance. Hyperparameter tuning is a crucial step to identify the best combination of model parameters that yield superior results. GridSearchCV was employed for XGBoost, and fine-tuning was carried out for the MLP model based on previous knowledge and experimentation.

3. **Model Training:**
   Each model was trained using the designated training set with the optimized hyperparameters. For XGBoost and TF-DF models, we leveraged their respective implementations in Scikit-learn and TensorFlow Decision Forests. For XGBoost and MLP models, we utilized the XGBoost and TensorFlow libraries.

4. **Model Evaluation:**
   To evaluate the predictive prowess of each model, we used the testing set, which was not seen during training or hyperparameter tuning. We calculated various performance metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R^2). These metrics provide comprehensive insights into the accuracy and robustness of the models.

By adhering to this comprehensive model development process, we ensured that each model was appropriately trained, optimized, and thoroughly evaluated. The models' performance metrics served as a reliable basis for selecting the most effective predictive model for ambulance demand. Detailed analysis and comparison of the models' performance will be discussed in subsequent sections of the study, enabling us to make informed decisions about resource allocation and enhance public health resource management.

3.4 Evaluation Metrics

The choice of evaluation metrics for the comparison of Random Forest, Artificial Neural Network, and XGBoost models in predicting the volume of ambulance dispatches was influenced by their relevance and suitability to the context.

The Mean Absolute Error (MAE) metric was adopted as it computes the mean of the absolute differences between predicted and actual values, independent of their direction (Willmott & Matsuura, 2005). This metric is crucial in the setting of ambulance demand prediction, offering an indication of the average prediction error in the number of calls, which is essential to ensure that the appropriate quantity of resources are deployed timely and efficiently.

The Root Mean Squared Error (RMSE) metric was selected, given its role in measuring the square root of the average of squared discrepancies between predicted and actual values (Hyndman & Koehler, 2006). This metric is valuable in the context of ambulance demand prediction as it indicates the efficacy of the model in terms of accurately forecasting the hourly call volume.

The Coefficient of Determination (R^2) metric was employed, given its capability to measure the proportion of variance in the dependent variable (ambulance calls) that can be explained by the independent variables (weather data, traffic data, population data, and public holiday data). This metric is pivotal in the setting of ambulance demand prediction, revealing the model's capacity to capture the relationships between the independent variables and the dependent variable.

In summary, the chosen evaluation metrics, due to their relevance and suitability, provide comprehensive insights into the accuracy, precision, and overall performance of the machine learning models in predicting ambulance demand. These insights further guide model selection and refinement to yield the most reliable and effective forecasting tool.

3.5 Model Interpretation and Discussion

The culminating phase of the methodology focuses on the interpretation of the outcomes derived from the machine learning models and a discussion of their implications, guided by previous research and models of interpretation (Goldstein et al., 2015). This encompasses pinpointing the model that provides the most accurate predictions, delving into the significance of the features used in the model, and contemplating how the model can be leveraged to enhance the management of ambulance resources in Dublin.

The model offering the most accurate predictions is identified by comparing the performance metrics of each model, namely the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2). The model yielding the lowest MAE and RMSE alongside the highest R^2 value, as proposed by Wickramasuriya et al. (2019), is considered the most proficient.

Subsequently, the importance of the features is examined based on the feature importance metrics from the most accurate model. This evaluation provides an understanding of the key factors influencing ambulance demand and can be used to make informed decisions about resource allocation.

Finally, the practical application of the model is considered. Specifically, how its predictive capabilities can inform the efficient and effective deployment of ambulance resources in Dublin. This application draws from the emergency service logistics literature, aligning predictive insights with strategic deployment (Andrienko et al., 2013). This may include the optimization of ambulance schedules, strategic positioning of ambulances based on predicted demand hotspots, and other proactive measures to ensure that ambulance services are always available when needed.

4 Design Specification

4.1 Architecture

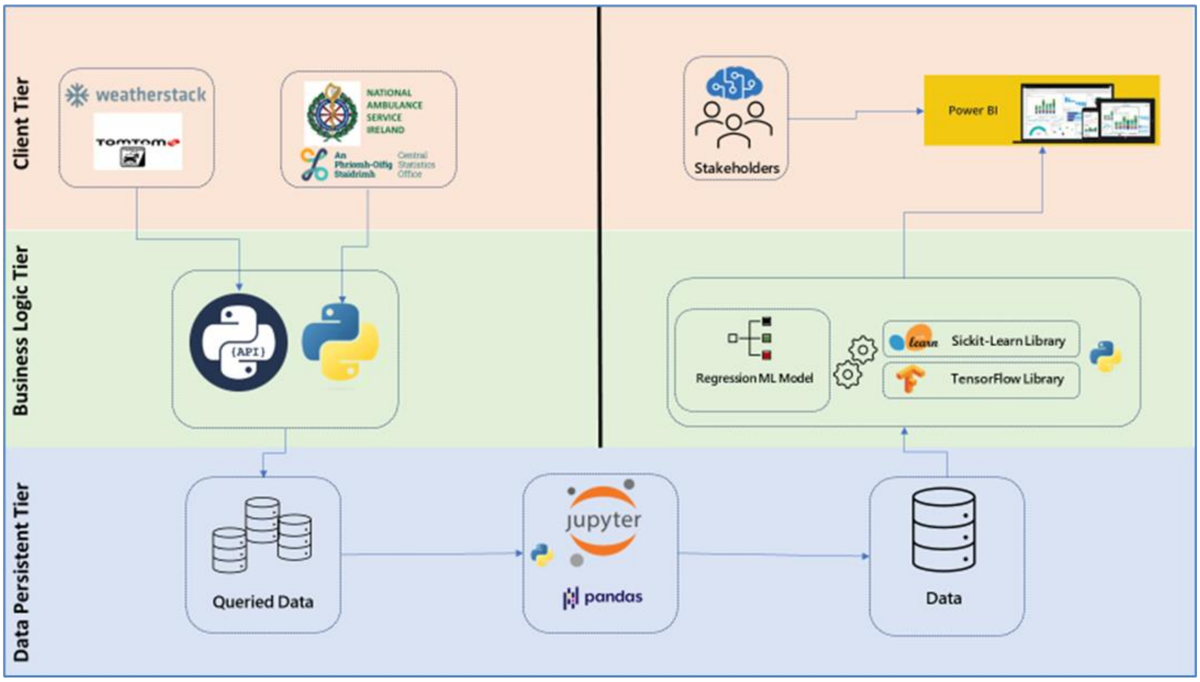

Figure 1: Ambulance Demand Prediction Multitier Architecture

As illustrated in Figure 1, the architecture for the Ambulance Demand Prediction system is structured into a multi-tier framework. The architecture comprises three core layers:

Data Persistence Layer: This foundational layer is tasked with the collection, storage, and pre-processing of various types of data. It handles data from diverse sources like the ambulance call-out data from the National Ambulance Service (NAS), meteorological data from Met Éireann, and public holiday data. The operations associated with data pre-processing, which encompass cleaning, merging, feature engineering, and scaling, are executed in this layer.

Business Logic Layer: This middle layer encapsulates the machine learning models employed for predicting ambulance demand, which include algorithms like Random Forest, XGBoost, and deep learning models such as the Multi-Layer Perceptron (MLP). This layer oversees the processes related to model training, hyperparameter tuning, and validation.

Client Tier: This uppermost layer serves as the user interface, facilitating the presentation of prediction results in an easily interpretable and accessible format. It may incorporate visualization tools to present model outcomes and predictions effectively.

This multi-tier architecture design is instrumental in ensuring efficient data management, robust training, and validation of predictive models, and a user-friendly presentation of prediction outcomes. Consequently, it offers a comprehensive and cohesive framework for accurate ambulance demand prediction.

4.2 Framework

The proposed solution to the research question regarding the forecast of ambulance demand will encompass the creation and assessment of machine learning techniques. Specifically, this undertaking aspires to evaluate the efficacy of diverse machine learning methodologies, comprising XGBoost, Random Forest, and Multilayer Perceptron (MLP) models, in forecasting ambulance necessities.

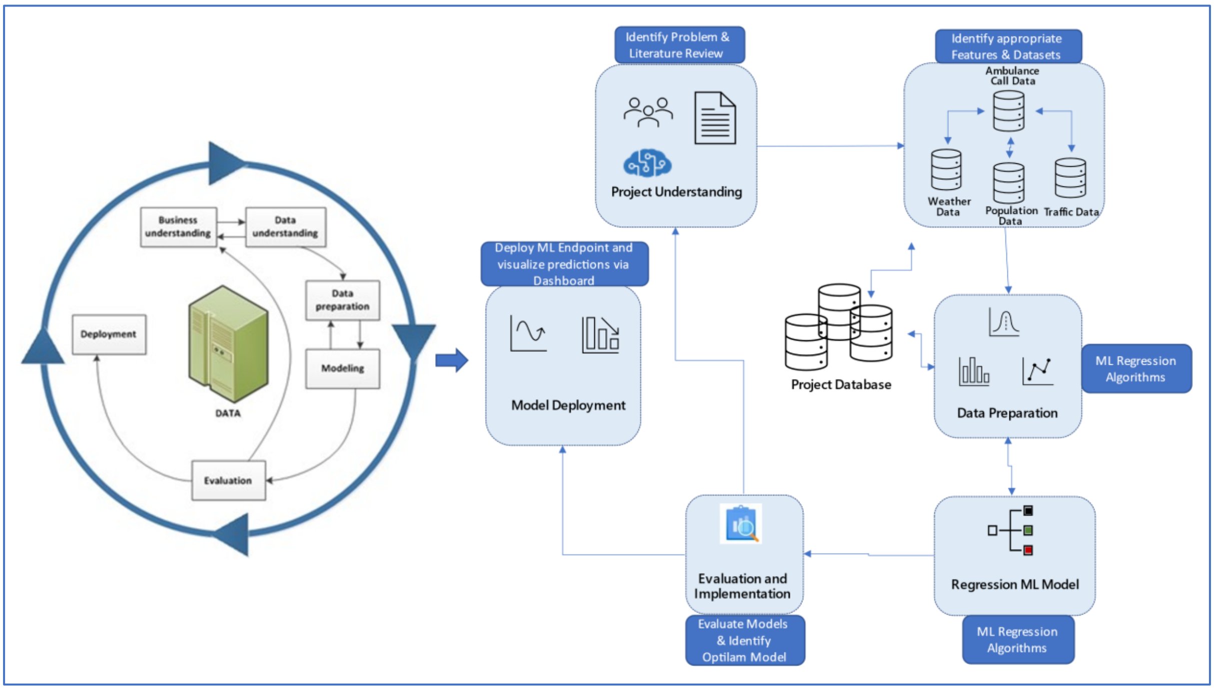

Figure 2. Ambulance Demand Prediction Methodology

In order to maintain an organized and methodical plan for the suggested research undertaking, a modified rendition of the Cross Industry Standard Process for Data Mining (CRISP-DM) approach will be employed, as depicted in Figure 2. The CRISP-DM approach furnishes an all-encompassing structure that supervises the entire data mining operation, starting from comprehending the issue to the final deployment of the model. The subsequent segment details the six phases of the CRISP-DM approach and the manner they will be utilized in this investigation.

Project Understanding: This phase encompasses the recognition of the issue and the delineation of the research inquiries to be examined. As articulated in the introductory section, the commercial challenge is the requisite for precise prediction of ambulance demand in Dublin, Ireland, with research questions concentrating on discerning the best machine learning techniques for forecasting ambulance demand and understanding how the engineering of features and the integration of novel features might augment the model's functionality.

Data Understanding: During this phase, pertinent data sources and variables will be pinpointed, and the data will be aggregated and processed. This includes obtaining publicly accessible ambulance call information from the National Ambulance Service (NAS) and correlated data sets, like public holidays and climatic conditions, to create an exhaustive dataset for examination.

Data Preparation: This phase entails purifying and altering the data to render it suitable for scrutiny. This will include cleansing the data, managing absent values, and selecting and engineering features. These data sources will then be amalgamated in readiness for the subsequent modeling phase. Exploratory data analysis (EDA) will be performed to uncover insights into the data and detect any outliers or inconsistencies.

Modelling: During this phase, the machine learning model will be constructed utilizing a variety of algorithms, encompassing neural networks, XGBoost, and random forests. The efficacy of each model will be appraised using fitting evaluation metrics, like MAE, RMSE, and R-squared. The model that excels in these metrics will be selected as the optimal one.

Evaluation: In this phase, the models will undergo an assessment to ascertain their capability in forecasting ambulance demand. The evaluation metrics chosen for this purpose are detailed in section 3.3. The outcomes will be juxtaposed with the initial research inquiries to verify that they have been suitably addressed.

Deployment: The concluding phase comprises incorporating the chosen model into a functional environment to make predictions on anticipated data, which will guide future demand forecasts. The model's applicability will be showcased on a PowerBI platform to demonstrate how it can be employed to aid decision-making in resource allocation and to augment the responsiveness of ambulance services.

Through adherence to the CRISP-DM approach, this study will guarantee an orderly and methodical method for solving the ambulance demand prediction challenge, commencing with comprehending the issue to the final deployment of the model. This adherence ensures that the research maintains rigor and comprehensiveness, and that the findings are dependable and valid.

4.3 Models and Algorithms

In pursuit of a solution to predict ambulance demand, this research project necessitates the construction and assessment of various machine learning models. The central goal is to compare the effectiveness of several machine learning algorithms, which encompass XGBoost, Random Forest, and Multilayer Perceptron (MLP) models, in accurately predicting ambulance demand.

XGBoost: The eXtreme Gradient Boosting (XGBoost) model is an implementation of the gradient boosting algorithm that has proven effective in numerous machine learning challenges. Its core advantage lies in its ability to handle both numerical and categorical data, its capability to manage missing values, and its flexibility in modeling complex, non-linear relationships, which are anticipated in the prediction of ambulance demand.

Random Forest: The Random Forest model is a well-known ensemble learning method that employs multiple decision trees and averages their output to enhance the prediction accuracy and prevent overfitting. Its strength lies in its ability to handle high-dimensional datasets and provide insight into feature importance, which will be beneficial in understanding the underlying factors influencing ambulance demand.

Multilayer Perceptron (MLP): The Multilayer Perceptron (MLP) is a type of artificial neural network that consists of multiple layers of neurons, including an input layer, one or more hidden layers, and an output layer. MLPs are versatile and capable of learning complex patterns in the data. They have been successfully applied to various tasks, including time-series data prediction. In this research, the MLP will be used to model the temporal patterns and relationships in the ambulance demand data, making it suitable for capturing non-linear dependencies over time.

Each of these models offers distinct advantages, and a comparative analysis of their performance on the ambulance demand prediction problem will yield valuable insights. The selection of the final model will be based on its performance metrics, its interpretability, and its ability to generalize well to unseen data.

5 Implementation

5.1 Data Transformation

The data transformation process was carried out meticulously to ensure the accuracy and effectiveness of the predictive models. It involved the following steps:

Data Collection: We collected relevant datasets, including ambulance call-out records, weather data, public holiday information, and other pertinent data sources.

Data Cleaning: To rectify any errors or inconsistencies in the data, data cleaning procedures were executed. This encompassed removing duplicates, handling missing or null values, and addressing outliers.

Data Aggregation: The main dataset was grouped by station area and merged with other relevant datasets based on the date and time of each ambulance call-out. The data was then aggregated into hourly intervals to provide a suitable level of granularity for analysis.

Feature Engineering: Categorical variables, such as call priority and station area, were transformed into numerical forms using one-hot encoding. Additionally, feature engineering was employed to include accumulated counts of call-outs and other potentially beneficial features. At detailed breakdown of the feature engineering steps carried out in this study is discussed in 5.1.1.

Data Scaling: To ensure the effectiveness of the machine learning models, numerical variables were scaled using StandardScaler to bring them to a similar scale.

5.1.1 Feature Engineering

Feature engineering is critical for improving the predictive capabilities of machine learning models. In this study, three distinct feature sets were drawn from datasets relating to ambulance demand in Dublin: the Basic, Full, and a Post XGBoost Feature Selection set.

Basic Feature Set: This dataset comprised readily available data from the Dublin ambulance dataset, such as 'Station Area', 'Date', 'Time Of Call', 'Count', 'day_of_week', 'month', 'weekend', and 'season'.

Full Feature Set: This dataset encompassed all features of the Basic set, enhanced with additional sourced data, including weather parameters ('rain', 'temp', 'wetb', etc.), time-series features ('Count for this hour in last x hours/days'), and the flag for public holidays.

Post XGBoost Feature Selection Set: After applying a XGBoost feature selection technique, we obtained a reduced set of features deemed most important for prediction: 'Count_in_last_7days', 'Count_in_last_14days', 'Count_in_last_28days', 'weekend', 'Date', 'Time Of Call', 'Count', 'Station Area', and 'Public Holiday'. This will be further covered in the Evaluation section **6.2**.

5.2 Tools and Languages

The implementation of the research project utilized Python as the primary programming language. Python's readability and versatility made it a suitable choice for this study. The following libraries were instrumental in data processing, manipulation, and machine learning modeling:

- **Pandas:** For data manipulation and pre-processing.
- **NumPy:** For numerical computations and operations on arrays.
- **TensorFlow:** For implementing neural networks and Decision Forests.
- **Scikit-learn:** For implementing machine learning models and evaluation metrics.

For handling more advanced models like XGBoost and Multilayer Perceptron (MLP), we harnessed the capabilities of the XGBoost and TensorFlow libraries. These libraries are renowned for their extensive functionalities, adaptability, and robustness in building machine learning models.

To facilitate data visualization and exploratory data analysis, we utilized the Matplotlib and Seaborn libraries. These libraries offer a wide range of functionalities, enabling comprehensive and insightful exploration of the data. The Integrated Development Environment (IDE) selected for this study was Google Colab, known for its interactive nature, user-friendly interface, and the added advantage of GPU usage, making it an ideal platform for executing complex machine learning algorithms and data analysis tasks..

5.3 Models Used

A variety of machine learning models were trained and evaluated to predict ambulance demand, as detailed in Section 4.3. The models included:

- Random Forest
- eXtreme Gradient Boosting (XGBoost)
- Multilayer Perceptron (MLP)

Each model was trained on 70% of the data in time-series order, with 20% reserved for testing and the remaining 10% for validation. Hyperparameter tuning was performed for each model to optimize their performance. The effectiveness of these models was evaluated based on relevant metrics, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2).

The time-series split is essential in this context, as ambulance demand is expected to have temporal patterns and dependencies that need to be considered in the model training and evaluation (Wickramasuriya et al. 2019). Using a time-based split ensures that the validation process accounts for the chronological order of the data, preserving the temporal structure, and providing a more realistic evaluation of the model's ability to generalize to unseen future data.

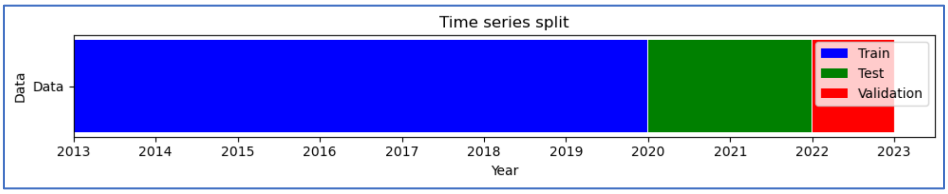

Figure 3. Time-Series Split

6 Evaluation

In this section, an exhaustive evaluation is presented of the machine learning models implemented for predicting ambulance demand. These models include the XGBoost, Random Forest as implemented through TensorFlow's Decision Forest, and the Multi-Layer Perceptron (MLP) Neural Network. To evaluate these models, we employed three key performance indicators: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination, R^2. In addition, the evaluation considers the impact of feature engineering on the models' predictive ability, reflecting our discussions on the importance of feature selection and the construction of new attributes based on the existing ones.

6.1 Overview of Model Performance

Table 2 below provides a comprehensive evaluation of the performance of the selected machine learning models across different feature sets. Distinct trends emerge when observing the evaluation metrics, indicating the individual models' sensitivity and adaptability to the feature variations.

Table 2: Performance Metrics for Different Models and Feature Sets

Model	Feature Set	R^2	% Change in R^2	RMSE	% Change in RMSE	MAE	% Change in MAE
MLP Neural Network	Basic Feature Set	0.085	-	0.901	-	0.516	-
	Full Feature Set	0.264	+209%	0.807	-10.3%	0.472	-8.5%
	Post XGBoost Feature Selection	0.299	+250%	0.788	-12.5%	0.465	-9.9%
XGBoost	Basic Feature Set	0.232	-	0.825	-	0.554	-
	Full Feature Set	**0.369**	+59.1%	**0.748**	-9.3%	**0.496**	-10.5%
	Post XGBoost Feature Selection	**0.367**	+58.2%	**0.749**	-9.2%	**0.497**	-10.3%
TensorFlow Decision Forest	Basic Feature Set	0.215	-	0.833	-	0.555	-
	Full Feature Set	0.368	+71.2%	**0.748**	-10.2%	**0.496**	-10.6%
	Post XGBoost Feature Selection	0.365	+69.8%	**0.750**	-10%	**0.496**	-10.6%

The MLP Neural Network, when examined, showcased noticeable improvements. Transitioning from the Basic Feature Set to the Full Feature Set resulted in a significant boost, with the R^2 value jumping by 209%. Furthermore, the Post XGBoost Feature Selection accentuated its performance, culminating in a 250% increase in R^2. This pattern suggests that the MLP Neural Network benefits immensely from feature enhancement and selection.

On the other hand, the XGBoost model, starting with a respectable R^2 of 0.232 on the Basic Feature Set, experienced a surge in its performance with the Full Feature Set, achieving a 59.1% improvement in R^2. Interestingly, even when the number of features was reduced in the Post XGBoost Feature Selection, the XGBoost model retained a comparable performance, achieving an R^2 of 0.367. This demonstrates the model's efficiency, as it was able to achieve near-identical results with fewer features, underlining its advantages in terms of computational efficiency and resistance to overfitting.

TensorFlow's Decision Forest showed consistent growth as it moved from the Basic Feature Set to the Full Feature Set, with its R^2 improving by 71.2%. The subsequent feature set, Post XGBoost Feature Selection, brought a minor decrease in the R^2, but the change was relatively insubstantial.

From the data in Table 2, it becomes clear that the incorporation of a comprehensive feature set plays a crucial role in enhancing a model's predictive capabilities. Across all models, the Full Feature Set consistently outperformed the Basic Feature Set. However, the outcomes from the Post XGBoost Feature Selection were varied, highlighting the challenges and complexities associated with feature selection.

Considering these observations, the marginally best performer was XGBoost model with the Full Feature Set, with the XGBoost feature set achieving almost identical performance. In addition, both models proved to maintain a high-performance level with fewer features offering a notable advantage. We will further probe into model performance in the subsequent section.

6.2 XGBoost

The XGBoost model, renowned for its gradient boosting framework, displayed significant performance, particularly when applied to a fully featured dataset and a dataset post-XGBoost feature selection. As discussed in the literature review, the success of XGBoost in this application could be attributed to its ability to capture complex non-linear relationships and its resistance to overfitting.

Table 3: XGBoost Model Performance

Evaluation Metric	Full Feature Set	Basic Feature Set	Post XGBoost Feature Selection
MAE	0.49673	0.554875	0.497342
RMSE	0.74803	0.825101	0.749228
R^2	0.369156	0.232465	0.367134

An integral aspect of the XGBoost model's operation is its feature importance mechanism. This mechanism rates the significance of each feature in the prediction model. The importance of features varies, with some contributing more to the predictions than others. The feature importance for our model is demonstrated in Figure 4 below:

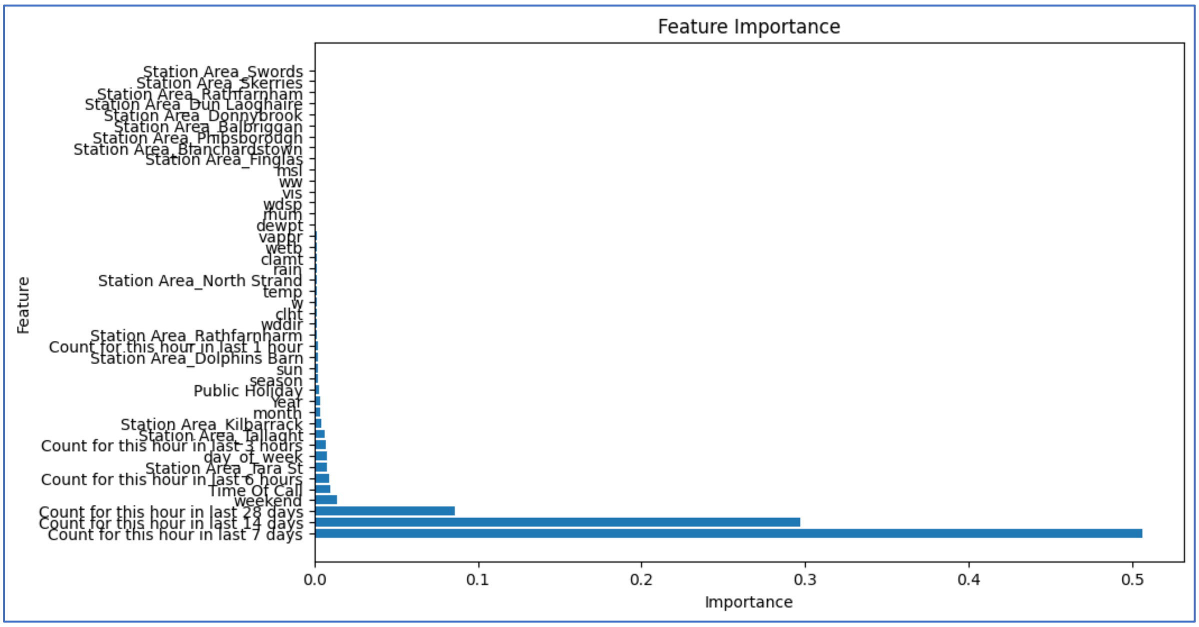

Figure 4. XGBoost Feature Importance

The feature importance chart for the model, as shown in Figure 4, offers significant insights. Notably, historical data features such as "Count for this hour in last 7 days", "Count for this hour in last 14 days", and "Count for this hour in last 28 days" score high, implying they play a critical role in predicting ambulance demand. This indicates the existence of temporal patterns in ambulance demand. Other influential features include "weekend", "Time Of Call", and "Station Area", which respectively denote the day of the week, the time the call was made, and the geographical station area.

Based on this feature importance information, a new dataset ‘Post XGBoost Feature Selection‘ was created containing only the most influential features. This smaller dataset focuses on the most critical elements, thereby reducing overfitting, improving prediction accuracy, and increasing computational efficiency.

Therefore, the XGBoost model, combined with the crucial step of feature selection informed by feature importance, provides a robust and efficient approach to predict ambulance demand.

6.3 Random Forest

The Random Forest model's effectiveness is reflected in its consistent performance across the different feature sets, as shown in Table 4. The Mean Absolute Error (MAE) values ranged from 0.496314 (Full Feature Set) to 0.555137 (Basic Feature Set), and marginally increased to

0.496880 when using the dataset post XGBoost feature selection. This suggests that the model can handle high dimensional data and deliver reliable predictions.

The Root Mean Squared Error (RMSE) values ranged from 0.748428 (Full Feature Set) to 0.833942 (Basic Feature Set), showing a decrease to 0.750034 with the dataset post XGBoost Feature Selection. The R^2 values were similar across all feature sets, implying that the proportion of the variance in the dependent variable that is predictable from the independent variables was consistent, regardless of the complexity of the feature set.

Table 4: Random Forest Model Performance

Evaluation Metric	Full Feature Set	Basic Feature Set	Post XGBoost Feature Selection
MAE	0.496314	0.555137	0.496880
RMSE	0.748428	0.833942	0.750034
R^2	0.368485	0.215929	0.365773

6.4 MLP Neural Network

The performance of the MLP Neural Network with different feature sets is detailed in Table 5. From the table, the MAE values for the full feature set, basic feature set, and post XGBoost Feature Selection were 0.472612, 0.516209, and 0.465147, respectively. Notably, the dataset post XGBoost Feature Selection exhibited the best performance with an MAE of 0.465147.

Regarding the RMSE values from Table 5, the full feature set had an RMSE of 0.807909, which went up to 0.900717 for the basic feature set, and then slightly decreased to 0.788218 after XGBoost feature selection. This trend reaffirms the significance of feature selection in bolstering the prediction accuracy of the model.

The R^2 scores, as shown in the table, were distinct for each feature set. The model trained on the Post XGBoost Feature Selection set had the highest R^2 score at 0.2996, suggesting it could explain approximately 29.96% of the variance in ambulance demand.

Table 5's data underlines the potential of deep learning models like the MLP Neural Network in effectively managing intricate prediction challenges such as ambulance demand forecasting, especially when they are tailored with the most suitable features.

Table 5: MLP Model Performance

Evaluation Metric	Full Feature Set	Basic Feature Set	Post XGBoost Feature Selection
MAE	0.472612	0.516209	0.465147
RMSE	0.807909	0.900717	0.788218
R^2	0.264117	0.085338	0.299551

6.5 Feature Engineering

Table 2 presented in section 6.1 illustrates the impact of feature engineering in the models' performance. Models trained on the Full Feature Set and Post XGBoost Feature Selection showed significant improvements over those trained with the Basic Feature Set.

The MLP Neural Network exhibited a substantial R^2 growth, moving from 0.085 to 0.264 and 0.299 with the Full Feature Set and Post XGBoost Feature Selection respectively. This underscores the value of comprehensive feature incorporation and methodical feature selection.

The significant role played by historical data ('Count for this hour in last x days') is evident in the feature importance given by the XGBoost model. These features top the list with a substantial margin, underlining their relevance in predicting ambulance demand. Features like 'weekend' and 'Time Of Call' also appear to add valuable information to the model.

These observations are consistent with prior literature, illuminating the factors driving ambulance demand in Dublin. Effective feature engineering not only amplifies predictive precision but also deepens our comprehension of the key drivers influencing ambulance demand, thereby facilitating better resource allocation.

6.6 Discussion

This study explores the application of three machine learning models - XGBoost, Random Forest, and MLP Neural Network - for predicting ambulance demand in Dublin city. Our goal was to leverage these sophisticated computational tools to bolster public health resource management. The results from our experiments contribute valuable insights regarding the utility of machine learning in public health settings.

With regards to performance, the XGBoost model demonstrated superior predictive ability, achieving an R-squared value of 0.37, RMSE of 0.75, and an MAE of 0.50 with the optimised feature set. This finding is consistent with prior research, which often highlights XGBoost's robustness in tackling a wide array of predictive tasks. Nonetheless, there were occurrences of both underestimation and overestimation of demand, indicating room for model enhancement through methods such as hyperparameter tuning or further refinement of the feature selection process.

The Random Forest model exhibited a level of effectiveness comparable to XGBoost, with identical performance metrics (R-squared value of 0.37, RMSE of 0.75, and an MAE of 0.50). This model's ability to handle nonlinear relationships and its robustness against overfitting likely contribute to its reliable performance.

The MLP Neural Network trained on the XGBoost feature set performed slightly below XGBoost and Random Forest. Despite achieving an R-squared value of 0.30, RMSE of 0.79, and an MAE of 0.47, this model may benefit from further refinement. Given the intricate nature of neural networks, adjustments such as adding additional layers, neurons, or modifying the activation function may enhance its predictive capability.

A comparison of these models underscores the unique strengths and weaknesses inherent to each model. Their applicability depends on the requirements and limitations of the specific use case. For example, while XGBoost boasts impressive performance, its intensive computational resource demand may be unsuitable for real-time or resource-constrained scenarios.

Feature engineering played a central role in improving model performance. Features such as "Count for this hour in last 7 days", "Count for this hour in last 14 days", "Count for this hour in last 28 days", and 'weekend' significantly influenced the predictive accuracy. However, the inclusion of certain features, including weather variables, seemed to introduce more noise than value to the models. This underlines the importance of careful feature selection and validation.

The study has certain limitations. One significant limitation is the static nature of the models, which contrasts with real-world situations that demand models capable of adapting in real-time to evolving data patterns. Future research could explore online learning models to address this limitation. Additionally, investigating other machine learning techniques, integrating additional data sources, and leveraging advanced feature engineering methods could further boost prediction accuracy.

In conclusion, this research reinforces the potential of machine learning for enhancing public health resource management, such as ambulance demand prediction. The results emphasize the necessity for judicious model selection, rigorous testing, and thoughtful feature engineering. This study serves as a foundation for future research aimed at refining and advancing this crucial area of study.

6.7 Limitations

While this study has yielded significant findings, it's important to acknowledge its limitations. Firstly, our data is limited to ambulance demand in Dublin, and while the models have proven effective in this context, they may not perform equally well in other regions. There may be unique factors at play in Dublin that the models have become overly fitted to, which could reduce their generalizability.

Additionally, the scope of the features considered, while broad, is not exhaustive. There may be other relevant features that we did not consider in our study, such as socio-economic factors, which could play a role in ambulance demand.

Finally, while our models performed well on our performance metrics, it's important to acknowledge the remaining error. The instances where the models overestimated or underestimated demand could represent situations where additional, unknown factors are influencing ambulance demand, indicating room for improvement in the models' predictive accuracy.

These limitations underscore the importance of continued research in this area. Future studies should explore these aspects further, using our research as a foundation upon which to build a more comprehensive and precise understanding of ambulance demand.

6.8 Overprediction vs. Underprediction

In predictive modeling, particularly in vital sectors like ambulance demand forecasting, it's crucial to navigate the balance between overprediction and underprediction, understand their implications, and devise strategies to mitigate associated risks.

When a model overpredicts, it anticipates a demand greater than the actual. While this ensures a heightened state of readiness with abundant resources in place, facilitating quick response to unforeseen demand spikes, it comes at a cost. Excessive resource allocation due to these overestimations can result in notable inefficiencies. Resources might be squandered, potentially increasing operational costs and diverting essential resources from other crucial areas.

On the contrary, underprediction leads the model to forecast a demand lower than the actual requirement. Though there might be short-term operational cost savings, as resources aren't over-allocated, the repercussions in a critical service context can be severe. Undervaluing demand can compromise emergency response times, risking lives. Additionally, consistently underestimating can overburden the available resources, possibly causing staff fatigue and a decline in service quality.

Given the gravity of ambulance services, a model veering towards overprediction might seem more prudent. Yet, it's imperative to assess the broader economic ramifications. The risk of underprediction, on the other hand, is grave since it can compromise patient safety. Striking the right balance is imperative. A potential strategy could be a dual-system approach, where a baseline resource allocation is guided by a model with a slight over predictive bias, augmented by a swift response mechanism to address unforeseen demand surges.

7 Conclusion and Future Work

In our pursuit of exploring and addressing the issues surrounding ambulance demand in Dublin, Ireland, this study has made notable strides in advancing the field of demand prediction and enhancing resource management within the emergency medical services (EMS) sphere. By comprehensively integrating machine learning techniques and evaluating numerous factors influencing demand, the research has not only effectively catered to Dublin's specific situation but also bridged a significant knowledge gap in the broader scientific literature.

7.1 Summary of Contributions

This study conducted an in-depth examination of ambulance demand in Dublin, Ireland, and developed a machine learning model to predict this demand, with the overarching aim of improving resource allocation within emergency medical services (EMS). The research has effectively filled the knowledge gap identified in the literature by developing an accurate prediction model specific to Dublin, while also enhancing our comprehension of the significant factors influencing ambulance demand.

Among the several machine learning algorithms explored, XGBoost proved to be the optimal model for predicting ambulance demand in Dublin. This aligns with the existing literature where studies like Ramgopal et al., (2021) highlighted the capabilities of XGBoost for accurate ambulance demand prediction. The robustness of XGBoost as a prediction model for ambulance demand is empirically validated through our findings.

Contrary to common assumptions, incorporating weather data into the model did not enhance its predictive power. This challenges the prevailing notion in the literature that weather significantly influences ambulance demand. Nonetheless, temporal factors, especially the introduction of accumulative counts, and societal factors like public events significantly improved the model's accuracy, reinforcing their importance in ambulance demand prediction.

7.2 Directions for Future Work

Considering the findings, several directions for future research can be suggested. Considering that weather data did not improve the model's prediction accuracy, it would be valuable to investigate other external factors that could potentially influence ambulance demand. Deepening our understanding of the relationship between different external factors and ambulance demand can lead to the creation of more accurate prediction models.

Moreover, this study focuses specifically on Dublin. The model's adaptability and reliability in different geographical contexts across Ireland are yet to be tested. Future studies could look at validating the model in these regions, adjusting for region-specific factors to ensure robust and reliable results.

In conclusion, this study has significantly contributed to the pressing issue of resource management in EMS by developing an accurate prediction model for ambulance demand. The findings provide important insights for Dublin's EMS and potentially for similar urban areas worldwide. Future work expanding on these findings has the potential to further enhance EMS resource management.

References

Andrienko, G. et al. (2007) “Geovisual analytics for spatial decision support: Setting the research agenda,” International Journal of Geographical Information Science, 21(8), pp. 839–857. Available at: https://doi.org/10.1080/13658810701349011.

Aringhieri, R. et al. (2017) “Emergency medical services and beyond: Addressing new challenges through a wide literature review,” Computers & Operations Research, 78, pp. 349–368. Available at: https://doi.org/10.1016/j.cor.2016.09.016.

Baker, J.R. and Fitzpatrick, K.E. (1986) “Determination of an optimal forecast model for ambulance demand using goal programming,” The Journal of the Operational Research Society, 37(11), p. 1047. Available at: https://doi.org/10.2307/2582689.

Chen, A.Y. et al. (2016) “Demand forecast using data analytics for the preallocation of ambulances,” IEEE Journal of Biomedical and Health Informatics, 20(4), pp. 1178–1187. Available at: https://doi.org/10.1109/jbhi.2015.2443799.

Goldstein, A. et al. (2015) “Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation,” Journal of Computational and Graphical Statistics, 24(1), pp. 44–65. Available at: https://doi.org/10.1080/10618600.2014.907095.

Hyndman, R.J. and Athanasopoulos, G. (2021) Forecasting: Principles and practice. Melbourne: OTexts.

Hyndman, R.J. and Koehler, A.B. (2006) “Another look at measures of forecast accuracy,” International Journal of Forecasting, 22(4), pp. 679–688. Available at: https://doi.org/10.1016/j.ijforecast.2006.03.001.

Jin, R. et al. (2021) ‘Predicting emergency medical service demand with bipartite graph Convolutional Networks’, IEEE Access, 9, pp. 9903–9915. doi:10.1109/access.2021.3050607.

Kyrkou, C. et al. (2023) “Machine learning for emergency management: A survey and future outlook,” Proceedings of the IEEE, 111(1), pp. 19–41. Available at: https://doi.org/10.1109/jproc.2022.3223186.

Larose, D.T. and Larose, C.D. (2015) Data Mining and Predictive Analytics. Somerset: Wiley.

Lin, A.X. et al. (2020) ‘Leveraging machine learning techniques and engineering of multi-nature features for National Daily Regional Ambulance Demand Prediction’, International Journal of Environmental Research and Public Health, 17(11), p. 4179. doi:10.3390/ijerph17114179.

Martin, R.J., Mousavi, R. and Saydam, C. (2021) “Predicting emergency medical service call demand: A modern spatiotemporal machine learning approach,” Operations Research for Health Care, 28, p. 100285. Available at: https://doi.org/10.1016/j.orhc.2021.100285.

McLay, L.A., Boone, E.L. and Brooks, J.P. (2012) “Analyzing the volume and nature of emergency medical calls during severe weather events using regression methodologies,” Socio-economic Planning Sciences, 46(1), pp. 55–66. Available at: https://doi.org/10.1016/j.seps.2011.09.002.

Met Éireann (2023) Historical Weather Data Ireland, Met Éireann - The Irish Meteorological Service. Available at: https://www.met.ie/climate/available-data/historical-data (Accessed: 03 August 2023).

National Ambulance Service Delivery Plan 2020 (no date). Available at: https://www.nationalambulanceservice.ie/publications/service-planning/nasdeliveryplan2020.pdf (Accessed: March 4, 2023).

Ramgopal, S. et al. (2019) "Weather and temporal factors associated with use of emergency medical services," Prehospital Emergency Care, 23(6), pp. 802–810. Available at: https://doi.org/10.1080/10903127.2019.1593563.

Ramgopal, S. et al. (2021) "Use of a metalearner to predict emergency medical services demand in an urban setting," Computer Methods and Programs in Biomedicine, 207, p. 106201. Available at: https://doi.org/10.1016/j.cmpb.2021.106201.

Setzler, H., Saydam, C. and Park, S. (2009) "EMS Call Volume Predictions: A comparative study," Computers & Operations Research, 36(6), pp. 1843–1851. Available at: https://doi.org/10.1016/j.cor.2008.05.010.

Smart Dublin (2023) Fire brigade and ambulance call outs DCC, Fire Brigade and Ambulance Call Outs DCC - Dataset. Available at: https://data.smartdublin.ie/dataset/fire-brigade-and-ambulance (Accessed: 03 August 2023).

Time&Date.com (2023) Ireland Public Holidays, Holidays and observances in Ireland in 2023. Available at: https://www.timeanddate.com/holidays/ireland/2023 (Accessed: 03 August 2023).

Vile, J.L. et al. (2016) "Time-dependent stochastic methods for managing and Scheduling Emergency Medical Services," Operations Research for Health Care, 8, pp. 42–52. Available at: https://doi.org/10.1016/j.orhc.2015.07.002.

Wickramasuriya, S.L., Athanasopoulos, G. and Hyndman, R.J. (2018) 'Optimal Forecast Reconciliation for hierarchical and grouped time series through Trace Minimization', Journal of the American Statistical Association, 114(526), pp. 804–819. doi:10.1080/01621459.2018.1448825.

Willmott, C.J. and Matsuura, K. (2005) "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," Climate Research, 30, pp. 79–82. Available at: https://doi.org/10.3354/cr030079.