An In-Depth Survey on Predictive Flight Delay Analysis, Flight Maintenance Strategies, and Elevating Customer Satisfaction in Aviation Operations

Supriya RK [1], Alaina Aanam [2], Kaavya SV [3], Anshika Kakani [4], Saranya K

[1,2,3,4,5]*Dayananda Sagar Academy of Technology and Management,Bangalore,India*
[1]supriya-ise@dsatm.edu.in,[2] alainaaanam@gmail.com, [3]svkaavya25@gmail.com,
[4]saranya11d12@gmail.com,[5]anshikakakani@gmail.com.

Abstract

This paper presents a comprehensive survey that meticulously explores the intricate domains of predictive flight delay analysis, aviation maintenance practices, and their profound impact on customer satisfaction. Delving into cutting-edge methodologies and technological trends, the study critically examines the multifaceted landscape of forecasting flight delays and optimizing maintenance protocols. Additionally, the research systematically identifies gaps in existing literature, offering a nuanced understanding of areas where further research and innovation are warranted. With a focus on data- driven approaches and innovative models, the survey aims to be a valuable resource for researchers, practitioners, and stakeholders in the aviation industry, providing crucial insights into overcoming challenges associated with flight delays and improving overall aviation efficiency. By synthesizing diverse findings, this work contributes significantly to academic research while offering practical implications for aviation professionals, guiding future endeavors to enhance operational efficiency and elevate customer satisfaction.

Keywords- *data-driven approaches, flight delays, aviation operations.*

I. INTRODUCTION

In the dynamic realm of aviation, our survey paper embarks on an extensive exploration, unraveling the complexities of predictive flight delay analysis, innovative flight maintenance strategies, and their profound implications for customer satisfaction. More than a mere documentation of existing knowledge, this endeavor seeks to empower and elevate the work of researchers, offering a comprehensive synthesis that not only informs but profoundly simplifies the trajectory of future investigations.

At the core of our investigation lies predictive flight delay analysis, a critical domain in an era characterized by the relentless surge in global air travel. Delving beyond the surface, we meticulously trace the evolutionary journey of methodologies, ranging from classical statistical models to the expansive array of machine learning algorithms. Our scrutiny extends to benchmark datasets, serving as signposts that ease the researcher's journey in navigating the labyrinth of predictive flight delay analysis. Navigating the technological landscape, our focus sharpens on themethods steering predictive flight delay analysis. Beyond algorithmic enumeration, we delve into the intricacies of time series analysis, decision trees, random forests, support vector machines (SVM), gradient boosting, and neural networks. In each exploration, we aim not only to elucidate but to simplify, providing a roadmap for researchers to traverse this landscape with confidence and clarity. Transitioning seamlessly, we venture into the core of aviation sustainability—aircraft maintenance. Here, we introduce an innovative fault prediction framework, a fusion of an auto-regressive moving average (ARMA) model with data- driven techniques. It's not just about fault prediction; it signifies a paradigm shift toward real-time monitoring, offering researchers a novel perspective for future investigations in predictive maintenance. Our exploration extends beyond technical intricacies to the financial implications of delays on flight maintenance.Dissecting the factors influencing delays at international airports, we unravel the intricate interplay between operational practices and financial outcomes.

This isn't just a documentation of challenges; it's an offering to researchers—a deep dive into the nexus of financial implications and operational strategies, simplifying the comprehension of this multifaceted

landscape. In the labyrinth of air travel's competitive landscape, our survey meticulously scrutinizes strategies for elevating customer satisfaction. By analyzing social media sentiments and dissecting passenger feedback, we unearth operational strategies aligned with passenger expectations. This isn't a mere pursuit of passenger contentment; it's a strategic analysis recognizing the symbiotic relationship between customer satisfaction and operational success. For researchers, this is an insight-rich territory, a goldmine of ideas for future studies on improving the passenger experience. A focal point in our narrative is the proposition of a novel fault prediction framework—an embodiment of innovation in predictive maintenance. Rooted in real-time monitoring, it transcends traditional life usage models, offering a lens into predictive accuracy through statistical features and temporal information. This isn't just a contribution; it's a robust foundation for researchers to build upon, providing a clear methodology for future investigations in fault prediction. As we navigate contemporary currents, our survey unveils prevailing industry trends. This segment serves as a reflection of current challenges and a clarion call for transformative solutions. Our exploration also delves into the challenges that permeate the aviation industry, offering researchers a nuanced understanding of the hurdles that demand innovative solutions.

Finally, we peer into the future trajectories of aviation research, inviting researchers to not merely observe but actively shape the

II. LITERATURE SURVEY

[1] This study addresses flight delays at John F. Kennedy International Airport by evaluating seven machine learning models. Notably, the Decision Tree model excelled with an accuracy of 0.9778, demonstrating superior performance in predicting flight delays. Tree-based ensemble classifiers, specifically Random Forest and Gradient Boosted Tree, also displayed strong predictive capabilities. In contrast, non-tree-based models like Logistic Regression, KNN, Gaussian Naïve Bayes, and SVM showed comparatively lower effectiveness. The dataset, covering flights from November 2019 to December 2020 and sourced from Kaggle, was processed using the "DEP_DELAY" variable for binary classification. Leveraging insights from prior studies, the research underscores the significance of machine learning models, particularly Decision Tree, in enhancing the prediction of flight delays. The adoption of weighted evaluation measures addresses dataset imbalances, contributing robust findings to the field.

[2] This research utilizes machine learning algorithms, including decision trees, random forests, SVM, gradient boosting, and neural networks, for flight delay prediction. Relevant features, such as departure/arrival time, airline, airport, and weather conditions, contribute to accurate predictions. Model evaluation includes metrics like accuracy, precision, recall, F1 score, and AUC-ROC. Drawing insights from studies like Addu et al., the proposed strategy covers data quality, feature selection, algorithm selection, model evaluation, and interpretability. Referenced studies address challenges in aviation data, real-time delay prediction, handling unbalanced data, and accurate forecasting. The comprehensive approach aims to leverage machine learning for enhanced understanding and prediction of flight delays in the aviation industry.

[3] This study conducted for a minor airline with approximately 10,000 flights annually and around 30 airplanes aims to provide results generalized for similarly sized airlines. The data, sourced from an internal database covering peak seasons from 2013 to 2019, was analyzed using modified IATA codes for delay causes. The methodology involved contingency table analysis and bivariate correspondence analysis, examining relationships between delay length, aircraft type, occupancy, and various factors. The study delves into financial implications, transparency advocacy, and statistical tests, including Pearson’s chi-square test and Cramer’s V. Key findings include operational success, identification of delay influencers, insights into load factors, flight types, and timing patterns. The study reveals the evolution of delay causes and dependencies across multiple variables. Expert validation enhances the research's credibility, providing actionable insights for strategic optimization and continued operational efficiency.

industry's future. This isn't just an academic endeavor; it's a dynamic call to action, beckoning researchers to join the vanguard of transformative studies that propel the aviation industry towards resilience, innovation, and unparalleled efficiency.

[4] This study addresses the surge in en-route delays, exemplified by a 105% increase in 2018, totaling 19.1 million minutes, reported by EUROCONTROL. With delay costs ranging from EUR 32 to EUR 80,270, the research proposes a novel approach, classifying air traffic scenarios into risk classes based on expected delay costs, integrating weather data and external events. Utilizing random forest classifiers and

following the CRISP-DM process, the study shifts from individual flight delay prediction to focus on meteorological impact and introduces "air traffic scenarios." Analysis uncovers insights into delay frequencies concerning weather and external events, addressing data imbalance through undersampling. Classification results show precision and recall values exceeding 80% for the high-cost Class-3 risk category. The study lays the groundwork for intelligently combining diverse air traffic and environmental data, with future research directions focusing on larger datasets, time series considerations, and airport-specific attributes. The findings offer valuable insights for enhancing air traffic control scenario classification.

[5] This paper introduces a model focused on searching and identifying the fastest flights between source and destination. Utilizing open source/public APIs, the model, implemented in Neo4j, converts data into JSON format and achieves an accuracy of 98.2% for delay prediction. The proposed flight delay prediction model employs clustering sampling, a random forest classifier, and a path-finding algorithm. Random forest is chosen based on information criteria, minimizing information loss for accurate prediction. Results indicate the model's superiority, with CRS departure time and scheduled departure time significantly impacting delays. Flight search results are demonstrated for Ahmedabad to Delhi, and delay prediction probabilities are calculated for better user experience. The conclusion highlights the model's effectiveness, outperforming related works, and suggests future enhancements using Graph Recurrent Neural Network for improved accuracy and additional features like flight booking.

[6] The paper addresses the issue of flight delays in civil aviation, emphasizing the substantial costs and impact on passenger satisfaction. It focuses on the predictive analysis of flight delays, particularly at Beijing International Airport (PEK), a data-driven and intelligent airport. PEK's predictive model combines flight, air traffic control, weather, and environmental data to locate and decompose the causes of delays in real-time. The study distinguishes between direct factors (e.g., weather) and indirect factors (e.g., historical delays) affecting flight delays. Machine learning and deep learning models, such as LSTM-AM (Long Short-Term Memory with Attention Mechanism), are explored for accurate predictions. The findings highlight the significance of considering both direct and indirect factors for effective delay

prediction, and the LSTM-AM model outperforms other algorithms. The research contributes to enhancing airport management, resource allocation, and passenger experience by providing actionable insights for mitigating delays. Future research may explore multi-airport interactions and consider the historical data of airport clusters.

[7] This paper presents a novel approach for flight delay prediction using a three-phase model: data collection and pre-processing, pre- training model with stack denoising autoencoders (SDA), and model optimization with the Levenberg-Marquart (LM) algorithm. The study systematically evaluates the impact of the number of denoising autoencoders and neurons on precision, accuracy, and processing time. The proposed model, SDA-LM, outperforms alternative structures, demonstrating significant improvement in accuracy, precision, sensitivity, and F1 measure, particularly when using a balanced dataset. The innovative combination of SDA and LM algorithm proves effective in handling noisy data, enhancing the model's predictive capabilities. The study showcases the potential of this methodology for accurate and reliable flight delay prediction, surpassing the performance of existing methods.

[8] This study addresses a critical gap in current airline maintenance systems by introducing a novel framework that combines auto-regressive moving average (ARMA) modeling with data-driven techniques for enhanced fault prediction and timely maintenance recommendations. The research demonstrates the superiority of the support vector regression model over traditional life usage models in predicting critical aircraft valve removals. The application of a generalized linear model also yields effective results, showcasing the framework's versatility. Through a comprehensive methodology and case study validation, the study advances fault prediction for complex engineering systems, particularly in the realm of aircraft maintenance. The results emphasize the framework's potential to significantly improve predictive accuracy, marking a noteworthy advancement in fault event prediction based on failure and inspection data.

[9] This study innovatively addresses the deficiency in current airline maintenance systems by introducing a pioneering framework that amalgamates auto-regressive moving average (ARMA) modeling with data-driven techniques, enhancing fault prediction and facilitating timely maintenance recommendations. Significantly, the research highlights the efficacy of the support vector regression model over traditional life

usage models in predicting critical aircraft valve removals. Additionally, the application of a generalized linear model demonstrates effective outcomes, underscoring the adaptability of the proposed framework. The study, validated through a robust methodology and a pertinent case study, propels the field of fault prediction in complex engineering systems, particularly within aircraft maintenance. The findings underscore the framework's substantial potential to markedly enhance predictive accuracy, marking a notable stride in forecasting fault events based on failure and inspection data.

[10] This study addresses the dynamic challenges of the airline industry by proposing a predictive system employing feed-forward neural networks and multivariate regression models. The developed framework accurately forecasts passenger demand, average fare, and no-show passengers, presenting a valuable tool for Origin Destination (OD) managers. Achieving high accuracies of 93-96%, the models significantly contribute to decision-making, workload reduction, and revenue optimization. The paper highlights the potential application of transfer learning for class expansion and suggests future research avenues, including the incorporation of additional predictors, optimization of forecasting models, and exploration of advanced AI methods to enhance adaptability in the ever-evolving airline industry.

[11] This article explores the application of machine learning (ML) in predictive maintenance for aircraft engines. ML and deep learning (DL) techniques, such as Random Forest (RF), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM), are examined for their effectiveness in regression and classification tasks. The study focuses on comparing classification and regression techniques, evaluating them using metrics like Root Mean Squared Error (RMSE) for regression and Accuracy, Precision, and Recall for classification. The dataset, sourced from NASA, consists of sensor measurements from degrading turbofan engines, contributing to the predictive maintenance analysis. Predictive maintenance is crucial for the aviation industry to enhance safety, reduce operational costs, and estimate the Remaining Useful Life (RUL) of aircraft engines.
Results indicate that LSTM, within the classification framework, performs optimally for predictive maintenance, achieving a high level of accuracy. On the regression side, Random Forest is identified as the best-performing algorithm. The study emphasizes the importance of selecting the appropriate technique based on research objectives, with classification techniques offering advantages in terms of accuracy and ease of implementation for predictive maintenance on aircraft engines.

[12] The paper proposes a predictive maintenance system for manufacturing units, utilizing machine learning techniques such as Random Forest Regression and LSTM. The primary objective is to minimize downtime and maintenance costs by predicting equipment failures based on historical sensor data. The proposed web application cleanses and analyzes data, identifies correlations, and trains suitable machine learning models. The use of LSTM is highlighted for its ability to handle sequential data effectively. The system aims to provide real-time updates to operators, enabling proactive maintenance and enhancing overall productivity. The study contributes to the growing emphasis on Industry 4.0 and smart systems, offering potential for sustainable manufacturing practices.

[13] The research methodology employed in this study comprises three essential components: establishing research questions, conducting a bibliometric analysis, and undertaking a comprehensive review to identify trends and research gaps. The review focuses on the period up to 2015, specifically evaluating state-of-the-art predictive maintenance (PdM) techniques. The research questions aim to identify key contributors, journals, organizations, and countries in the field, primary applications in
the aircraft industry, and challenges and opportunities facilitated by new technologies. The search parameters include aerospace- benchmark datasets and highlights the prevalence of time series data in aircraft maintenance, with the Commercial Modular Aero- Propulsion System Simulation (C-MAPSS) being a frequently used benchmark. While existing reviews touch on various aspects of PdM, this paper distinguishes itself by being the first exhaustive review solely focused on aircraft, promising to identify challenges and opportunities comprehensively. The need for a consolidated review dedicated to aircraft systems, considering diverse datasets and applications, becomes evident, fostering a better understanding of PdM advancements in the aerospace industry.

[14] This research introduces a hybrid sentiment analysis approach using classical ML algorithms and a Deep Neural Network (DNN) applied to TripAdvisor hotel reviews. Classical ML algorithms exhibit limited accuracy (<70%), while the proposed DNN consistently achieves >98% accuracy, revealing a significant performance gap. The study emphasizes the effectiveness of hybrid models and underscores the potential of

DNNs in sentiment analysis. The research identifies a performance gap in classical ML algorithms, addressed by the superior accuracy of the DNN model. The conclusion highlights the success of the combined approach in sentiment analysis and underscores the importance of advanced techniques like deep learning in overcoming limitations of traditional ML methods. The study suggests future exploration of varied coding models and datasets for comprehensive sentiment analysis.

[15] This research delves into sentiment analysis of air travel- related tweets from major airlines, employing a unique hybrid approach that blends classical Machine Learning (ML) algorithms

Figure 1. Sentiment Analysis Software.

focused papers published since 2015, emphasizing well-cited works. The bibliometric analysis examines authors, journals, and countries, revealing China and the USA as prolific contributors, particularly through IEEE journals. The paper introduces with advanced Deep Neural Networks (DNN). The study showcases the effectiveness of SVM, ANN, and CNN for sentiment classification, with CNN demonstrating superior performance. Additionally, association rule mining uncovers influential factors such as flight delays and in-flight comfort. The findings underscore the significance of a holistic approach beyond sentiment analysis, emphasizing the need to comprehend underlying factors for an enhanced customer experience in the airline industry. The study acknowledges limitations, suggesting future exploration of diverse language datasets to broaden its scope.

[16] The paper explores the impact of perceived safety on customer satisfaction in air travel, considering the purpose of travel (business or pleasure). It identifies that safety plays a crucial role in shaping passenger perceptions. Business travelers, being more rational and less price-sensitive, exhibit a lower sensitivity to safety concerns compared to pleasure travelers. For pleasure travelers, an increase in perceived safety positively influences satisfaction and customer loyalty. Safety is suggested to be treated as an asset, with proactive communication of high safety standards. The study incorporates various components like ground handling and the flight experience as drivers of overall customer satisfaction. The findings imply that airlines should strategically manage safety perceptions to enhance customer satisfaction and loyalty. This research provides actionable insights for airline managers to target marketing efforts based on the distinct needs of business and pleasure travelers.

III. CONCLUSION

In conclusion, our survey paper delves into the intricate realms of predictive flight delay analysis, innovative flight maintenance approaches, and the pivotal domain of customer satisfaction within the aviation industry. The amalgamation of diverse findings not only significantly contributes to our understanding but also lays the groundwork for practical applications, fostering operational efficiency and customer contentment. The examination of fault prediction frameworks, coupled with an in-depth analysis of the financial repercussions of delays, provides a holistic perspective on the challenges faced by the aviation sector. Simultaneously, our investigation into strategies for elevating customer satisfaction offers actionable insights for industry stakeholders and aviation professionals. As we navigate through the prevailing industry trends and challenges, it becomes apparent that there are avenues yet unexplored. Future research endeavours could focus on refining machine learning models,

exploring the integration of varied datasets, and gaining a more comprehensive understanding of evolving customer sentiments in the age of pervasive social media. This survey paper stands not merely as a compendium of existing knowledge but as an invitation to researchers to embark on a journey towards transformative studies. The identified gaps invite further exploration, promising a future where the aviation industry achieves new heights of efficiency and ensures a superior experience for every traveller.

REFERENCES

[1]Yuemin Tang. 2022. Airline Flight Delay Prediction Using Machine Learning Models. In Proceedings of the 2021 5th International Conference on E-Business and Internet (ICEBI '21). Association for Computing Machinery, New York, NY, USA, 151–154. https://doi.org/10.1145/3497701.3497725.

[2]S. Sharan, M. Sriniketh, H. Vardhan and D. Jayanth, "State-Of-Art Machine Learning Techniques to Predict Airlines Delay," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), Bengaluru, India, 2021, pp. 1-6, doi: 10.1109/FABS52071.2021.9702590.

[3]Zámková, Martina & Rojík, Stanislav & Prokop, Martin & Stolín, Radek. (2022). Factors Affecting the International Flight Delays and Their Impact on Airline Operation and Management and Passenger Compensations Fees in Air Transport Industry: Case Study of a Selected Airlines in Europe. Sustainability.
14. 14763. 10.3390/su142214763.

[4]Bardach, Markus & Gringinger, Eduard & Schrefl, Michael & Schütz, Christoph. (2020). Predicting Flight Delay Risk Using a Random Forest Classifier Based on Air Traffic Scenarios and Environmental Conditions. 1-8. 10.1109/DASC50938.2020.9256474.

[5]Z. Shu, "Analysis of Flight Delay and Cancellation Prediction Based on Machine Learning Models," 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2021, pp. 260-267, doi: 10.1109/MLBDBI54094.2021.00056.

[6]Wang, Fujun & Bi, Jun & Xie, Dongfan & Zhao, Xiaomei. (2022). Flight delay forecasting and analysis of direct and indirect factors. IET Intelligent Transport Systems. 16. 10.1049/itr2.12183.

[7]Yazdi, M.F., Kamel, S.R., Chabok, S.J.M. *et al.* Flight delay prediction based on deep learning and Levenberg-Marquart algorithm. *J Big Data* 7, 106 (2020). https://doi.org/10.1186/s40537-020-00380-z.

[8]Baptista, Márcia & Sankararaman, Shankar & de Medeiros, Ivo & Nascimento Jr, Cairo & Prendinger, Helmut & Henriques, Elsa. (2017). Forecasting Fault Events for Predictive Maintenance using Data-driven Techniques and ARMA Modeling. Computers & Industrial Engineering. 115. 10.1016/j.cie.2017.10.033.

[9]de Pater, Ingeborg & Reijns, Arthur & Mitici, Mihaela, 2022. "Alarm-based predictive maintenance scheduling for aircraft engines with imperfect Remaining Useful Life prognostics," Reliability Engineering and System Safety, Elsevier, vol. 221(C).

[10] P. H. K Tissera, A. N. M. R. S. P. llwana, K. T. Waduge, M. A. l. Perera, D.
P. Nawinna and D. Kasthurirathna, "Predictive Analytics Platform for Airline Industry," 2020 2nd International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 2020, pp. 108-113, doi: 10.1109/ICAC51239.2020.9357244.

[11] Azyus, Adryan. (2022). Determining the Method of Predictive Maintenance for Aircraft Engine Using Machine Learning. Journal of Computer Science and Technology Studies. 4. 01-06. 10.32996/jcsts.2022.4.1.1.

[12] Kane, Archit & Kore, Ashutosh & Khandale, Advait & Nigade, Sarish & Joshi, Pranjali. (2022). Predictive Maintenance using Machine Learning. 10.48550/arXiv.2205.09402.

[13] Stanton, I, Munir, K, Ikram, A, El-Bakry, M. Predictive maintenance analytics and implementation for aircraft: Challenges and opportunities.
Systems Engineering. 2023; 26: 216–237. https://doi.org/10.1002/sys.21651

[14] S. Gadri, S. O. Mehieddine, K. Herizi and S. Chabira, "An Efficient System to Predict Customers' Satisfaction on Touristic Services Using ML and DL
Approaches," 2021 22nd International Arab Conference on Information Technology (ACIT), Muscat,Oman, 2021, pp. 1-8, doi: 10.1109/ACIT53391.2021.9677167

[15].Kumar, S., Zymbler, M. A machine learning approach to analyze customersatisfaction from airline tweets. *J Big Data* 6, 62 (2019). https://doi.org/10.1186/s40537-019-0224-1.

[16] Ringle, Christian & Sarstedt, Marko & Zimmermann, Lorenz. (2011). Customer Satisfaction with Commercial Airlines: The Role of Perceived Safety and Purpose of Travel. Journal of Marketing Theory and Practice. 19. 459-472. 10.2753/MTP1069-6679190407.