Deep Learning Approaches for Analysis and Prediction of Flight Data

Gökhan GÜLŞEN
Necmettin Erbakan University

Mehmet Emin ÇETİN
Necmettin Erbakan University

To Cite This Chapter

Gülşen, G., & Çetin, M. E. (2024). Deep Learning Approaches for Analysis and Prediction of Flight Data. In S. Koçer & Ö. Dündar (Eds.), *Intelligent Systems and Optimization in Engineering* (pp. 1-8). ISRES Publishing.

Introduction

The aviation industry has grown rapidly in recent years, dramatically increasing the amount of passengers and cargo transported around the world. The increasing number of aircraft in the air necessitates the automation of airspace surveillance. It is well recognized that aircraft with varying purposes operate in distinct flight patterns. This growth has brought with it the need for greater efficiency, safety and cost optimization for companies and airports in the sector. The total cost of all disruptions in commercial aviation in the United States in 2008 was estimated at 33.5 billion dollars (Khaksar, H., et. al.). Only 79% of flights arrived on schedule in 2019, according to airline delay statistics, which resulted in tens of billions of dollars in damages, including lost demand, passenger charges, and other indirect costs (Kim, Y., J., Choi, S., Briceno, S., Mavris, D., 2016). In European airspace, a record was set on June 28, 2019 with 37288 flights(Nats, 2023). Improving factors such as flight safety, fuel consumption, operational costs and passenger satisfaction depends on effective analysis of flight data and accurate prediction of future flight conditions. In this context, issues such as flight route and flight delay prediction are of strategic importance for the aviation industry.

Today, flight route prediction is essential in terms of both safety and fuel consumption. Especially on routes with heavy air traffic, accurate route prediction can improve flight safety and reduce fuel consumption through optimal route selection. At the same time, accurate flight route prediction provides a significant advantage for airport congestion management and air traffic control.

On the other hand, flight delay forecasting is of great importance in terms of passenger satisfaction and operational costs. Delays cause serious financial losses for airlines and cause customer dissatisfaction by negatively affecting passengers' travel plans. Therefore, predicting potential delays using historical flight data allows airlines to plan more effectively and provide accurate information to customers.

The purpose of this study is to make flight route and flight delay predictions using deep learning methods. In particular, models that can process time series data such as CNN (LeCun, L., Bengio, Y., 1995) and LSTM (Hochreiter, S., Schmidhuber, J., 1997) will be used to analyze and predict flight data. The studies to be conducted in this context will contribute to increase the efficiency of flight operations and provide useful information for various stakeholders in the aviation industry.

Problem

This research study seeks to address two questions:

- Flight Path and Safety ADS-B (Automatic Dependent Surveillance-Broadcast) data(Faa, 2023) is very important for accurately tracking and predicting flight paths, as it contains instantaneous position, altitude, speed and direction information of aircraft. Especially in heavily trafficked airspaces, accurate data is needed for aircraft to navigate at a safe distance and determine their routes. This data allows the route to be predicted and optimized according to the air traffic density. By analyzing ADS-B data, air traffic management and airspace safety can be improved.
- It allows the observation of momentary changes in the flight process and the analysis of historical flight data to detect flight delays and ensure passenger satisfaction. Analyzing the causes of delays can improve passenger satisfaction and help regulate airport operations by predicting these problems in advance. Using flight data, delays due to weather, traffic density or technical reasons can be predicted. These predictions will increase the effectiveness of airlines in informing passengers and offering alternative solutions.

Machine Learning

A wide range of algorithms that may “learn” from data on their own are included in machine learning (ML) (Lindholm, A., Wahlstrom, N., Lindsten, F., Schon, T., B., 2022). Machine learning (ML) is defined as a type of data analysis where an algorithm learns on its own to extract information and generate predictions from a dataset. The training set is the name given to this collection of data. A separate data set, known as the test set, that does not intersect with the training set is then used to evaluate the algorithm. The data, the mathematical model, and the learning algorithm are the three main parts of a machine learning algorithm. Using the data, a machine learning system optimizes the parameters in the mathematical mode.

Artificial Neural Networks

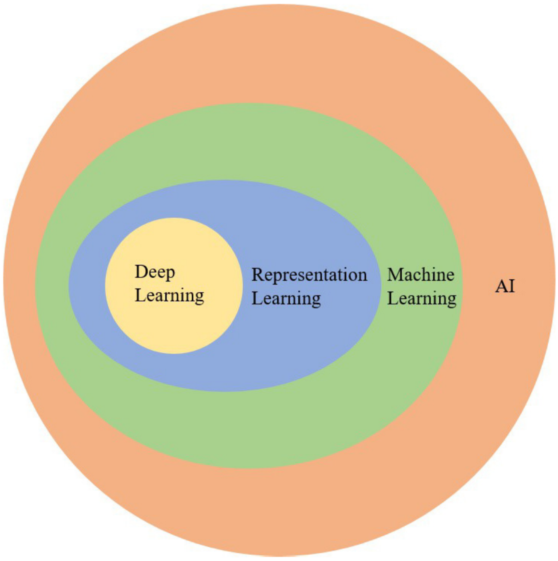
Using the human brain as an example, mathematical modeling of the learning process has led to the development of artificial neural networks (ANN). It replicates the architecture of the brain’s biological neural networks as well as their capacity for memory, learning, and generalization. Examples are used to carry out the learning process in artificial neural networks. Rules are established and input and output data are provided during the learning process.

Deep Learning

Deep learning models have demonstrated significant effectiveness across various domains. Examples include applications like speech recognition, language translation, and computer vision. Recent technological advancements, including self-driving cars and intelligent personal assistants, can be largely attributed to developments in deep learning techniques. Improved processing hardware and easier access to data are responsible for the rise in popularity of deep learning techniques in recent years. Goodfellow et al. (Goodfellow, I., Bengio, Y., Courville, A., 2016) illustrate deep learning using the Venn diagram presented in Figure 2. Deep learning models typically necessitate substantial datasets to exceed the performance of traditional machine learning approaches and are more computationally intensive.

Figure 1

Venn diagram illustrating that deep learning represents only a subset of artificial intelligence. (Goodfellow, I., Bengio, Y., Courville, A., 2016)



Whether to ML or DL depends mainly on the type of problem, the volume of the data set and the complexity requirements of the model. When deciding which is more suitable for developing prediction models for the problems in our thesis, we need to consider the following factors:

- Machine Learning (ML) methods generally give good results with smaller data sets and a limited number of features. ML methods are usually sufficient, especially when the data has numeric and categorical/class features, but does not contain a large number of samples.
- Deep Learning (DL) methods need large data sets. Large volumes of data, more processing power, and more time are typically needed for deep learning models. If your data set is large and contains more complex features (e.g. text data, images, sounds), then deep learning models can perform better.

Because of the large and our dataset's complexity (extreme location information, various delay reasons), Deep Learning (DL) methods were used.

Deep Learning Methods

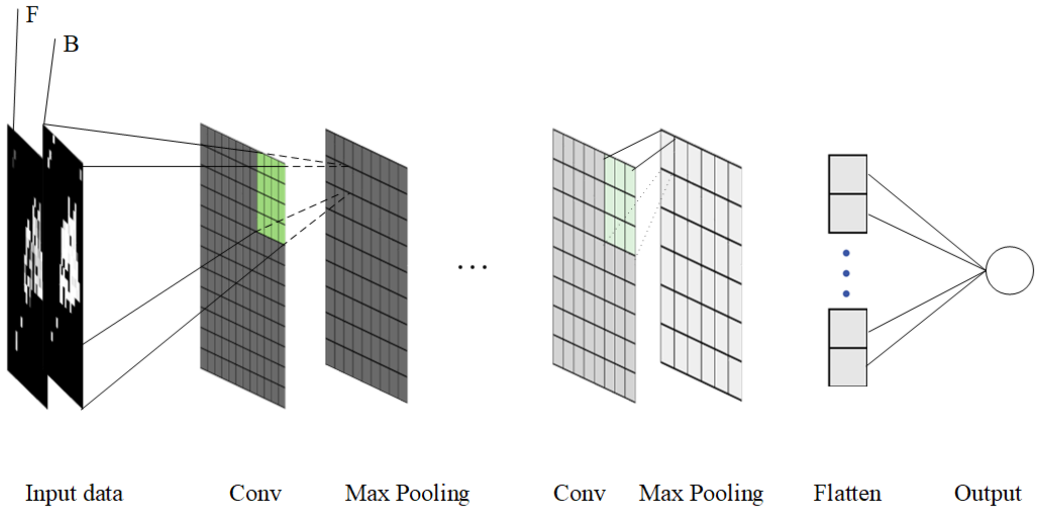
A subset of the larger machine learning family, deep learning focuses on learning data representations as opposed to task-specific methods. Unlike traditional machine learning methods, deep learning has the advantage of constructing deeper architectures that allow for learning more abstract information. Deep learning's capacity to automatically learn feature representations, doing away with the need for laborious manual feature engineering, is one of its primary characteristics. Deep learning has found practical applications in various fields(Coşkun, M., Yıldırım, Ö., Uçar, A., Demır, Y., 2017) through algorithms such as CNN, RNN and LSTM, each of which will be discussed in detail in the following sections.

What is CNN (Convolutional Neural Network)?

CNN (LeCun, Y., Bengio, Y., 1995) is a deep learning algorithm which can recognize local features in two- or three-dimensional data such as images, audio and time series. CNNs are widely used, especially in areas like natural language processing and image processing. Basically, it extracts certain patterns or convolution layers are used to extract features from the input data. and preserves important information by reducing the dimensions through pooling in Figure 2.

Figure 2

Structure of the CNN Model (Qiang L., Xinjia G. & Liu J)

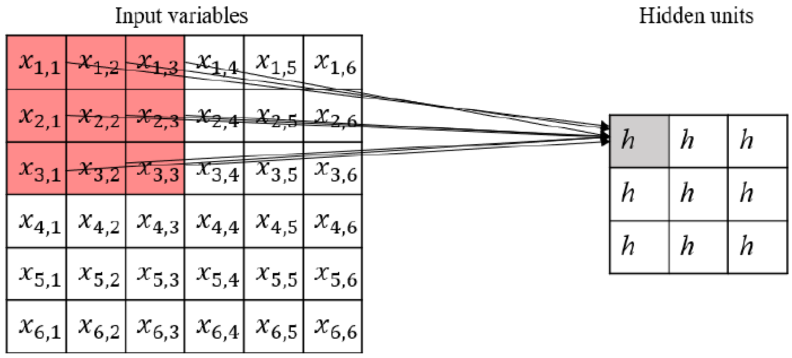


The main advantage of CNN is that it provides high accuracy by examining regionally significant parts of the data, not the whole data. Especially in-flight forecasting, CNN can detect short-term correlations and thus quickly learn important features in time series data.

*Convolution layers: The basic component of the CNN model. It processes the data by dividing it into small pieces through filters, see Figure 3 for an example. Each filter can be used to estimate 2D location data (e.g. latitude-longitude pairs), and convolution layers can be used to learn the environmental relationships and local dependencies of the data. This structure is particularly useful for predicting future position from past movements of position data.

*Pooling: A method used by CNN to decrease the size of the data being processed and escalate the model's processing speed. Pooling preserves important information by averaging or maximizing small regions of the data. In this way, the model operates more quickly and has less complexity.

Figure 3
Diagram of sparse interacting convolutional layers (Goodfellow, I., Bengio, Y., Courville, A., 2016)



CNN's Superiority and Advantages

Ability to Learn Local Dependencies: CNNs are successful in capturing local patterns in data thanks to the filters in their layers. In flight route and delay data, local and short-term changes in features such as speed, position, altitude are important for future predictions. CNN's ability to capture these short-term correlations improves its prediction performance.

Learning Long-Term Patterns in Time Series: Since flight data consists of time series data, CNN layers can provide accuracy in predictions by recognizing long-term patterns in these series. The convolution process of the filters is good at capturing changes in direction or changes in speed of the flight over time, thus learning important information when making flight route and delay predictions.

Faster Processing by Reducing Complexity: CNN models reduce the data size by filtering out unnecessary information thanks to the "pooling" process used to reduce the data size. This feature both increases processing speed and enables analysis with a smaller data size. This speed advantage provides a significant advantage in time-sensitive processes such as flight prediction.

Potential for Visual Data: Flight prediction can use not only textual or numerical data, but also maps or radar images that are 2D/3D visualizations of the flight path. CNN is ideal for working with this type of visual data. Especially in cases where the flight route needs to be followed visually, CNN models can give more successful results.

No Need for Data Processing to Reduce User Errors: CNN has the ability to learn without the need for many data pre-processing steps, especially thanks to its ability to work with raw data. This allows flight route and delay prediction data to be used in the model quickly and effectively.

Advantages that Differentiate CNN from Other Models

While algorithms such as RNN (Recurrent Neural Network) and LSTM are successful in learning long-term dependencies, they do not have the advantage of quickly learning local features and short-term patterns that CNNs provide. CNNs are good at identifying sudden changes in flight routing and short-term anomalies that cause delays.

Traditional methods such as Decision Trees or Regression Models may fail to capture long-term and short-term dependencies in the data at the same time. For this reason, CNN is observed to offer higher accuracy in flight prediction than traditional methods.

What is RNN (Recurrent Neural Network)?

One kind of neural network intended for processing sequential data is called a recurrent neural network (RNN) (Rumelhart, D., E., Hinton, G., E., Williams, R., J., 1986). Text, audio, and time series are examples of temporally oriented data that they can examine. By passing a secret state from one time step to the next, RNNs are able to accomplish this. Based on the input and the previous hidden state, the hidden state is updated at each time step. In sequential data, RNNs are good at identifying short-term dependencies, but they have trouble identifying long-term ones.

Inadequacy of RNNs

Neural networks called RNNs are made to handle sequential data and are good at recognizing short-term dependencies. However, they encounter an issue referred to as the vanishing gradient problem. Because of this issue, the gradients in the network get smaller and eventually vanish as the number of layers rises. RNNs are therefore unable to capture long-term relationships, thus losing important information from the past. In flight route and delay prediction, the long-term effects of historical data such as the previous position, speed, delay times and weather for a given flight need to be taken into account. Since RNNs cannot preserve this information, the accuracy of flight predictions decreases.

What is LSTM (Long Short Term Memory)?

One kind of RNN (recurrent neural network) that may store long-term dependencies in sequential input is called an LSTM (long short-term memory). Sequential data, including time series, text, and voice, may be processed and analyzed using LSTMs. They circumvent the issue of vanishing slope that plagues conventional RNNs by controlling the information flow via a memory cell and gates, allowing them to selectively store or discard information as needed. LSTMs are extensively employed in many different applications, including time series prediction, speech recognition, and natural language processing.

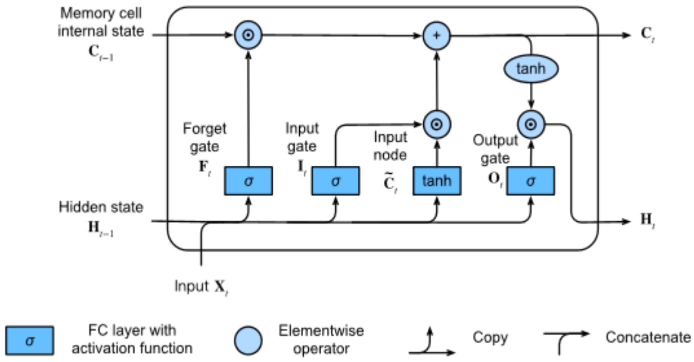
LSTM's Superiority

LSTM (Long Short-Term Memory) networks (Hochreiter, S., Schmidhuber, J., 1997) have the capacity to acquire long-term relationships and store important information in data. This is achieved through cell state and gate mechanisms:

- Input Gate: Determines how and to what extent new information is added to the cell.
- Forget Gate: Unnecessary information is removed from the cell, freeing up memory space.
- Output Gate: Controls which part of the information stored in the cell is added to the output.

Thanks to these structures, shown in Figure 4, LSTM is able to recall information from past time periods and make more accurate predictions by preserving long-term relationships.

Figure 4
Structure of the LSTM Model (Zhang A., Lipton Z., Li M. & Smola A.)



Conclusion

In this study, the challenges of flight route prediction and flight delay prediction are addressed as critical problems in the aviation industry. Flight route prediction plays a vital role in ensuring airspace safety and achieving fuel efficiency, while flight delay prediction is essential for improving operational efficiency and passenger satisfaction. To tackle these issues, real-time flight data is used to develop prediction algorithms with deep learning models.

The models can be employed in this study include CNN and LSTM architectures:

The LSTM model can be selected for its ability to process ttime-series data, which records the temporal relationships between past flight metrics such position, velocity, and route to make accurate predictions. The CNN model, on the other hand, is utilized for its capacity to extract spatial relationships and feature patterns within the dataset, providing strong results in predicting both flight routes and delay types. Deep learning methods are preferred over traditional machine learning algoritms because of their superior ability to handle large and complex datasets. Deep learning models, in contrast to conventional methods, do not require considerable human feature engineering since they automatically discover features from raw data. Moreover, deep learning excels at identifying intricate relationships and uncovering hidden patterns in high-dimensional data, making it particularly effective for datasets like ADS-B, which contain diverse and high-volume information.

Experimental evaluations demonstrate that the proposed prediction models successfully address the problems of flight route and delay prediction. The models consistently achieve high levels of accuracy and show their applicability in the aviation industry. These findings emphasize the potential of deep learning to enhance flight safety, reduce operational costs, and improve passenger satisfaction.

In conclusion, this study highlights that analyzing ADS-B data with deep learning methods provides effective solutions to operational challenges in the aviation industry. By leveraging the advanced capabilities of CNN and LSTM architectures, this research showcases how modern predictive models outperform traditional methods in terms of scalability, accuracy, and adaptability. Future work focuses on expanding the diversity of data sources and employing more advanced architectures to further improve model accuracy and address a broader range of aviation-related challenges.

References

Khaksar, H., & Sheikholeslami, A. (2019). Airline delay prediction by machine learning algorithms. *Scientia Iranica*, *26*(5), 2689-2702.

Akenine-Möller, T., & Ström, J. (2003). Graphics for the masses: a hardware rasterization architecture for mobile phones. In *ACM SIGGRAPH 2003 Papers* (pp. 801-808).

Kim, Y. J., Choi, S., Briceno, S., & Mavris, D. (2016, September). A deep learning approach to flight delay prediction. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)* (pp. 1-6). IEEE.

Nats, The european airspace challenge. https://nats.aero/blog/2019/09/ the-european-airspace-challenge/. Accessed: 2023-02-28.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995.

Li, Q., Guan, X., & Liu, J. (2023). A CNN-LSTM framework for flight delay prediction. *Expert Systems with Applications*, *227*, 120287.

Faa, Automatic dependent surveillance-broadcast (ADS-B). https://www.faa.gov/air_traffic/technology/adsb. Accessed: 2023-05-02.

Bts, https://www.bts.gov/

Lindholm, A., Wahlström, N., Lindsten, F., & Schön, T. B. (2022). *Machine learning: a first course for engineers and scientists*. Cambridge University Press.

Bengio, Y., Goodfellow, I., & Courville, A. (2017). *Deep learning* (Vol. 1). Cambridge, MA, USA: MIT press.

Coşkun, M., Yıldırım, Ö., Uçar, A., & Demır, Y. (2017). An overview of popular deep learning methods. *European Journal of Technique (EJT)*, *7*(2), 165-176.

LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, *3361*(10), 1995.

Pavlo, A. Lecture# 03: Data Formats & Encoding II.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Hochreiter, S. (1997). Long Short-term Memory. *Neural Computation MIT-Press*.

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. *arXiv preprint arXiv:2106.11342*.

About the Authors

Gökhan GÜLŞEN, MSc, is a professional with a background in software development. He earned his undergraduate degree from the National Defense University in Istanbul in 2015. He holds two master's degrees, one from Aksaray University (2021) and another from Necmettin Erbakan University (2024). He has work experience in various roles, starting as an Aircraft Maintenance Officer in the Turkish Air Force (2015-2020), followed by positions as a Software Development.
E-mail: gokhangulsen@ogr.erbakan.edu.tr, **ORCID:** 0000-0003-1876-7499

Mehmet Emin ÇETİN, Assoc. Prof., completed his BSc degree at Kocaeli University, Kocaeli in 2010 and MSc degree at Karadeniz Technical University, Trabzon in 2012. He received his PhD degree in Mechanical Engineering from Karadeniz Technical University, Trabzon, Turkey in 2017. His research interests are artificial intelligence, additive manufacturing, adhesives, composite materials and finite element analysis.
E-mail: mecetin@erbakan.edu.tr , **ORCID:** 0000-0002-6314-5261

Similarity Index
The similarity index obtained from the plagiarism software for this book chapter is 14%.