CHAPTER – 2

LITERATURE SURVEY

2.1 INTRODUCTION

With the rise in the airline industry and construction of new airports and runways, air traffic has increased dramatically in the past few years. It has put additional pressure on air traffic control system that handles thousands of flights per day. To avoid delays and collisions, air traffic management has to work efficiently. They face many problems like weather, communication, frequency congestion, work pressure, noise etc. They need to be aware that what is going on in their space as well as other sectors around them. If a problem arises, they need to act on it at the very moment. It puts a lot of pressure on air traffic controllers. A bad weather means a bad day for them. A controller needs to master the flow of communication on that frequency. Air traffic control is managing 50,000 flights daily, so they have to handle many flights at the same time. It means checking the altitude so that they don't collide against. Plus, multiple flights pass in the same route and during such situation if a flight face problem, they need to give special attention. For such reason several research papers that includes speech synthesis and speech recognition, language modelling, error categories and incorporation of data mining and neural networks in air traffic services, were referred and surveyed which were more relevant and much similar to our idea and the key points of those papers were noted for the implementation of our project.

2.2 RESEARCH PAPERS

Below are the research papers, which were surveyed for gaining information about existing research works which are related to our project. By surveying these below mentioned literature survey papers, it helps to build knowledge in our field of working, Data mining in Aviation. The abstracts of the literature survey papers includes speech synthesis and speech recognition, language modelling, error categories and incorporation of data mining and neural networks in air traffic services.

2.2.1 Using LSTM Encoder-Decoder Algorithm for Detecting Anomalous ADS-B Messages

ADS-B (Automatic Dependence Surveillance - Broadcast) plays a major role in the safe navigation of airplanes and air traffic control (ATC) management. But there is a lack of security mechanisms in this system and hence it is exposable to many attacks. Although many solutions were proposed, there is still a need of additional participating nodes (or sensors) which helps in verifying the location of the airplane by analyzing its physical signal. But most airplanes fly with ADS-B deployed in it so, applying those modifications is a bit complicated. [1]So an alternative security solution for detecting anomalous ADS-B messages is proposed which aims at the detection of spoofed or modified fake ADS- B messages sent by an attacker. The LSTM encoder-decoder algorithm is modelled here for safeguarding ADS-B messages from imposters. Using this model, aircraft can autonomously evaluate received ADS-B messages, thus helps to identify deviations from the legitimate flight path. The major drawback is, no recovery, in case of an internal system failure.

2.2.2 Emotional Statistical Parametric Speech Synthesis Using LSTM-RNNs:

This paper studies the methods for emotional statistical parametric speech synthesis (SPSS) using recurrent neural networks (RNN) with long short-term memory (LSTM) units. Two modeling approaches, i.e., emotion-dependent modeling and unified modeling with emotion codes, are implemented and compared by experiments. In the first approach, LSTM-RNN based acoustic models are built separately for each emotion type. A speaker-independent acoustic model estimated using the speech data from multi-speakers is adopted to initialize the emotion-dependent LSTM-RNNS. Inspired by the speaker code techniques developed for speech recognition and speech synthesis, the second approach builds a unified LSTM-RNN-based acoustic model using the training data of a variety of emotion types. In the unified LSTM-RNN model, an emotion code vector is input to all model layers to indicate the emotion characteristics of current utterance. Experimental results on an emotional speech synthesis database with four emotion types (neutral style, happiness, anger, and sadness) show that both approaches achieve significant better naturalness of synthetic speech than HMM-based emotion dependent modeling.[2] The emotion-dependent modeling approach outperforms the unified modeling approach and the HMM-based emotion-dependent modeling in terms of the subjective emotion classification rates for synthetic speech. Furthermore, the emotion codes used by the unified modeling approach are capable of controlling the emotion type and intensity of synthetic speech effectively by interpolating and extrapolating the codes in the training set. The LSTM encoder-decoder algorithm is modelled here for safeguarding ADS-B messages from imposters. Using this model, aircraft can autonomously evaluate received ADS-B messages, thus helps to identify deviations from the legitimate flight path.

2.2.3 Application of Data Mining in Air Traffic Forecasting

Many study centres has developed a model for the purpose of Air Traffic forecasting by using off-the-shelf data mining and machine learning techniques. Recent developments use data mining algorithms to predict the likelihood of previously un-connected airport-pairs being connected in the future, and the likelihood of connected airport-pairs becoming un-connected. [3]Despite the innovation of this research, it does not focus on improving the FAA's existing methodology for forecasting future Air Traffic levels on existing routes, which is based on relatively simple regression and growth models. So different approaches are investigated for improving and developing new features within the existing data mining applications in Air Traffic forecasting. As part of future work, machine learning techniques such as clustering and neural networks are getting applied to improve this model's performance.

2.2.4 Intelligent Air Traffic Control using Neural Networks

The air traffic control systems is again one of the most complex jobs thus has increased its complexity due to increase in aircrafts and airports. The air traffic controllers are responsible for taking complex decisions such as take-off, landing etc. The controllers depend on various parameters such as availability of runway, climate conditions, and other meteorological parameters and make the decisions which are much complicated and are highly prone to errors. Safety has to be considered when it comes to automation. This project deals with automation of existing air traffic control system using neural networks which comes under artificial intelligence. This has been found to be effective in many fields for making important and complex decisions. Back propagation network algorithm is used here for decision making. [4]The network is trained using some predetermined inputs and later the network will be capable of making decisions of its own with minimal or zero error. But Only Ideal Conditions were considered for evaluating and simulating the correctness of the systems, which are not applicable in real-time.

2.2.5 Using Recurrent Neural Networks for Slot Filling in Spoken Language Understanding

Semantic slot filling is one of the most challenging problems in spoken language understanding (SLU). In this paper, we propose to use recurrent neural networks (RNNs) for this task, and present several novel architectures designed to efficiently model past and future temporal dependencies. Specifically, we implemented and compared several important RNN architectures, including Elman, Jordan, and hybrid variants. [5]To facilitate reproducibility, we implemented these networks with the publicly available Theano neural network toolkit and completed experiments on the well-known airline travel information system (ATIS) benchmark. In addition, we compared the approaches on two custom SLU data sets from the entertainment and movies domains. Our results show that the RNN-based models outperform the conditional random field (CRF) baseline by 2% in absolute error reduction on the ATIS benchmark. We improve the state-of-the-art by 0.5% in the Entertainment domain.

2.2.6 Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition

Long Short-Term Memory (LSTM) is a recurrent neural network (RNN) architecture that has been designed to address the vanishing and exploding gradient problems of conventional RNNs. Unlike feedforward neural networks, RNNs have cyclic connections making them powerful for modeling sequences. They have been successfully used for sequence labeling and sequence prediction tasks, such as handwriting recognition, language modeling, and phonetic labeling of acoustic frames. The second approach builds a unified LSTM-RNN-based acoustic model using the training data of a variety of emotion types. However, in contrast to the deep neural networks, the use of RNNs in speech recognition has been limited to phone recognition in small scale tasks. [6]In this paper, we present novel LSTM based RNN architectures which make more effective use of model parameters to train acoustic models for large vocabulary speech recognition. We train and compare LSTM, RNN and DNN models at various numbers of parameters and configurations. We show that LSTM models converge quickly and give state of the art speech recognition performance for relatively small sized models.

2.2.7 TTS Synthesis with Bidirectional LSTM based Recurrent Neural Networks

Deep neural networks (DNN)-based text-to speech (TTS) systems have been recently shown to outperform decision-tree clustered context-dependent Hidden Markov Models (HMM) TTS systems. In this paper, Recurrent Neural

Networks (RNNs) with Bidirectional Long Short Term Memory (BLSTM) cells are adopted to capture the correlation or co-occurrence information between any two instants in a speech utterance for parametric TTS synthesis. [7] Experimental results show that a hybrid system of DNN and BLSTM-RNN, i.e., lower hidden layers with a feed-forward structure which is cascaded with upper hidden layers with a bidirectional RNN structure of LSTM, can outperform either the conventional, decision tree-based HMM, or a DNN TTS system, both objectively and subjectively. The speech trajectory generated by the BLSTM-RNN TTS is fairly smooth and no dynamic constraints are needed.

2.2.8 Context Dependent Recurrent Neural Network Language Model:

Recurrent neural network language models (RNNLMs) have recently demonstrated state-of-the-art performance across a variety of tasks. In this paper, we improve their performance by providing a contextual real-valued input vector in association with each word. This vector is used to convey contextual information about the sentence being modeled. By performing Latent Dirichlet Allocation using a block of preceding text, we achieve a topic-conditioned RNNLM. [8]This approach has the key advantage of avoiding the data fragmentation associated with building multiple topic models on different data subsets. We report perplexity results on the Penn Treebank data, where we achieve a new state-of-the-art. We further apply the model to the Wall Street Journal speech recognition task, where we observe improvements in word-error-rate.

2.2.9 LSTM Neural Networks for Language Modelling

Neural networks have become increasingly popular for the task of language modeling. Whereas feed-forward networks only exploit a fixed context length to predict the next word of a sequence, conceptually, standard recurrent neural networks can take into account all of the predecessor words. On the other hand, it is well known that recurrent networks are difficult to train and therefore are unlikely to show the full potential of recurrent models. [9]These problems are addressed by the Long Short-Term Memory neural network architecture. In this work, we analyze this type of network on an English and a large French language modeling task. Experiments show improvements of about 8 % relative in perplexity over standard recurrent neural network LMs. In addition, we gain considerable improvements in WER on top of a state-of-the-art speech recognition system.

2.2.10 Air Traffic Volume and Air Traffic Control Human Errors

Abstract Navigable airspaces are becoming more crowded with increasing air traffic, and the number of accidents caused by human errors is increasing. Based on many interviews, study, research and surveys of ATC safety experts, the ATC errors are categorized into three major categories. They are communication error, procedure error, and instruction error. The main objective of this paper is to evaluate the relationship between air traffic volume and human error in air traffic control (ATC). First, the paper identifies categories and elements of ATC human error through a review of existing literature, and a study through interviews and surveys of ATC safety experts. [10]And then the paper presents the results of an experiment conducted on 52 air traffic controllers sampled from the Korean ATC organization to find out if there is any relationship between traffic volume and air traffic controller human errors. An analysis of the experiment clearly showed that several types of ATC human error are influenced by traffic volume. We hope that the paper will make its contribution to aviation safety by providing a realistic basis for securing proper manpower and facility in accordance with the level of air traffic volume.

2.2.11 Data mining for air traffic flow forecasting: a hybrid model of neural network and statistical analysis

The objective of this paper is to build a hybrid model of neural network and statistical analysis of air traffic flow. The air traffic flow prediction plays a key role in the airspace simulation model and air traffic flow management system. In China the air traffic information in each regional control center has not integrated together by now. The information only in a single regional control center cannot reach the requirement of the current method based on 4-dimensional trajectory prediction. The new method is needed to solve this problem. Large collection of radar data is stored. But there is no effort made to extract useful information from the database to help in the estimation. [11]Data mining is the process of extracting patterns as well as predicting previously unknown trends from large quantities of data. This paper employs neural networks combined with the statistical analysis of historical data to forecast the traffic flow. Two models with different types and input data are proposed. The accuracy of two models is tested and compared to each other using flow data at an arrival fix in Beijing control center. This has been found to be effective in many fields for making important and complex decisions. Back propagation network algorithm is used here for decision making. The result shows that these models are feasible for practical implementations. The suitable models for different prediction conditions are also suggested.

2.2.12 Speech recognition using Recurrent Neural Prediction Model

The neural prediction model (NPM) proposed by Iso and Watanabe is a successful example of a speech recognition neural network with a high recognition rate. This model uses multilayer perceptrons for pattern prediction (not for pattern recognition), and achieves a recognition rate as high as 99.8% for speaker-independent isolated words. [12]This paper proposes a recurrent neural prediction model (RNPM), and a recurrent network architecture for this model. The proposed model very significantly reduces the size of the network, with as high a recognition rate as the original model, and with a high efficiency of learning, for speaker-independent isolated words.

2.3 SUMMARY

Various dimensions and techniques of Air Traffic Control System were briefed in the survey paper. Above list of works were surveyed for several factors with respect to our proposed idea. The factors include processing and time efficiency along with the system processing. The efficiency of the model depends on how well it is used by the controllers and how far it will be useful for them to control the airflow on non-critical situations and ideal conditions. However, all these proposals have some disadvantages. No recovery in case of internal system failure. Noise is again a pretty concerning factor, which could leave to misinformation. The most common demerit in all these papers is, they don’t actually focus a bit much on the stress faced by the controllers, which could lead to human errors. So our model mainly focuses on this specific factor called “Human stress reduction” and automation is applied here to reduce their work pressure.