Chapter 11
Deep Learning for Feature Engineering–Based Improved Weather Prediction:
A Predictive Modeling

Partha Sarathi Mishra
https://orcid.org/0000-0002-5129-3339
North Orissa University, India

Debabrata Nandi
North Orissa University, India

ABSTRACT

Weather prediction has gained a point of attraction for many researchers of variant research communities. The emerging deep learning techniques have motivated many researches to explore hidden hierarchical pattern in the great mass of weather dataset for weather prediction. In this chapter, four different categories of computationally efficient deep learning models—CNN, LSTM, CNN-LSTM, and ConvLSTM—have been critically examined for improved weather prediction. Here, emphasis has been given on supervised learning techniques for model development by considering the importance of feature engineering. Feature engineering plays a vital role in reducing dimension, decreasing model complexity as well as handling the noise and corrupted data. Using daily maximum temperature, this chapter investigates the performance of different deep learning models for improved predictions. The results obtained from different experiments conducted ensures that the feature engineering based deep learning study for the purpose of predictive modeling using time series data is really an encouraging approach.

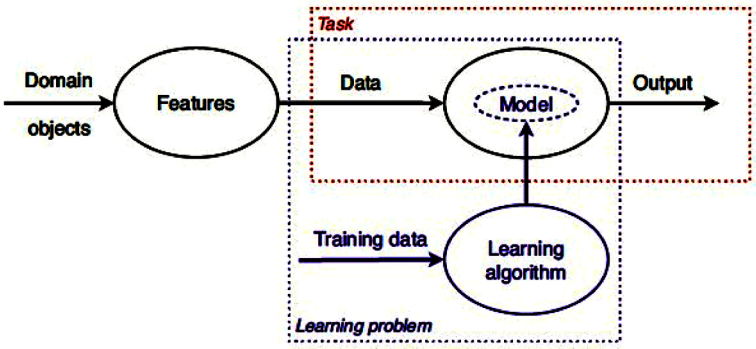
DOI: 10.4018/978-1-7998-6659-6.ch011

INTRODUCTION

The task of classification in pattern recognition (Dehuri et al. (2011), data mining (Panda et al. (2015)), and big data analytics (Acharjya et al. (2015); Mishra et al. (2016)) has gained attention of many researchers for developing a robust and accurate model to predict unseen data. In all the above spheres, the accuracy of the model depends on the quality of data being used. In every view of the machine learning techniques, some input data is supplied to get certain output. The input data supplied involves features, which are in the form of structured columns. These features with certain specific features are generally asked by the algorithms to run more efficiently. Hence, the feature engineering plays a crucial role in the machine learning environment. In current scenario, many of the data, scientists spend 60% of their time in cleaning and organizing data. As per survey in Forbes, the rest of the time is spent in building training sets (3%), a dataset collection (19%), and mining data for patterns (9%), Refining algorithms (4%), and others (5%). The missing value is one of the most common problems that we come across while preparing the data for efficient modeling using machine learning algorithms. It affects the machine learning models to a great extent. Here, where the feature engineering stands.

Machine learning is the organized study of algorithms and systems that increase their knowledge or performance with experience. A large number of classification methods (Wu, 2007; Bennette, 2014; Dehuri & Cho, 2010) as well as prediction methods (Mishra & Dehuri, 2012; 2014, 2016; Mishra, 2018) have already been developed by different researchers as mentioned. Figure 1 below describes how the machine learning algorithm works for a given address problem. Here, the model addresses the task; learning algorithms to solve the problem which leads to the development process. In summary, we can say that machine learning deal with the use of an exact feature in the process of building an exact model to accomplish exact tasks.

Figure 1. Machine learning model using features



Weather prediction has gained a point of attraction for many researchers of variant research communities. The emerging deep learning techniques have motivated many researches to explore hidden hierarchical pattern in the great mass of weather dataset for weather prediction. It is very important, because it decides the future climate change of a locality and provides a baseline for publishing the new findings on environmental principles and technology. Many of our daily activities and businesses also depend on weather conditions. Unpredictable weather conditions may also lead to hazards as well as property loss.