Basically the division of two classes in 2D or multidimensional is done by a hyper plane. This results in separation of two different classes is given in Figure 6.

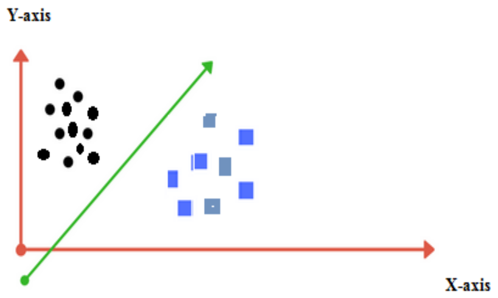

Figure 6. Separation in less complex data

It is very difficult to achieve the division if the data is quite. So complete separation of two classes in the x-y plane is done by the transformation and addition of one more plane i.e. z –axis to divide the plane into classes. Here transformations are basically the kernels. These separations are made so as to divide similar data sets at particular plane as shown in Figure 7. In SVM, having a linear hyper-plane between these two classes is simple. Do we really need to manually add this function to have a hyper-plane? No, SVM has a kernel trick method. These are the processes that take poor-dimensional input space and turn it into a stronger-dimensional space, i.e. transform non-separable case into separable aspect and these processes are kernels. It is mostly helpful in the problem of non-linear separation. Just put, it does some highly complicated information transformations, then figure out how to separate the information based on the labels or outputs you have described.

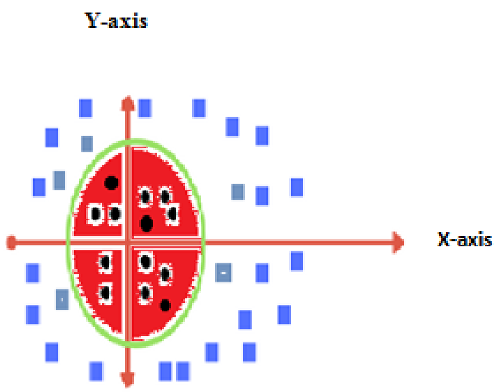

Figure 7. Separation in Complex Data

Prediction for the input vector (I_{new}) is mentioned below in equation 3 which is generated through the input vector (I_q) and each support vector (qi) by taking the dot product of both i.e.

$$Inew = Iq.\ qi \quad (2)$$

$$f(x) = \mathrm{C}(0) + \mathrm{SUM}(ri * Inew) \quad (3)$$

here coefficients C(0) and ri estimated by the training data and thus the above mentioned equation imply calculation for the inner products of recent input vector (I_q) including all the support vectors in the available training dataset.

Linear SVC (Support Vector Classifier) main objective is to return the best fit hyper plane in division of the classes. Using certain other features of python we can predict the values quite easily like matplotlib etc for data visualisation.

4. Results

Both the methodologies are applied on the same data set "weather archive jena 2009-2016"(420528 tupples) from kaggle datasets which includes attributes in terms of seconds but the after conversion it into hourly basis the overall tupples that are trained becomes 1751. The predicted result is quite different in terms RMSE value. The error rate in LSTM is 0.427 while in case of SVM it is 0.768. Figure 8 shows the graphs in that are obtained while the model implementation.

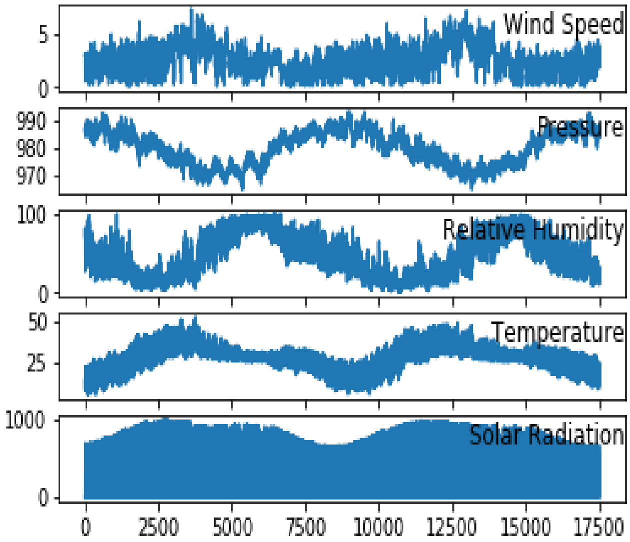

Figure 8. Visualized Parameters

These are the visualized parameters i.e. the parameters that were taken under consideration while the implementation of both the methodologies it includes parameters (wind speed, pressure, relative humidity, temperature, solar radiation).

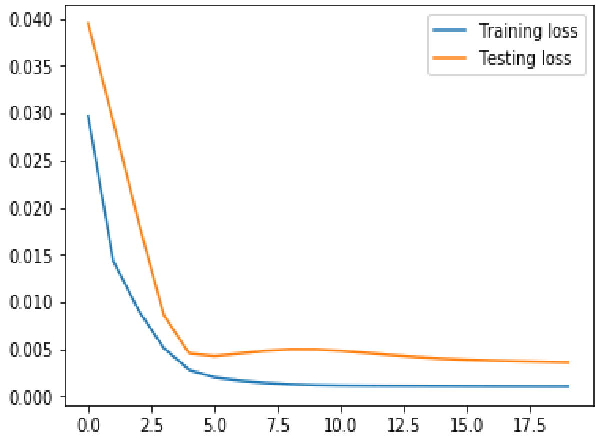

Figure 9. Training and Testing Loss

Figure 9 given above shows the graph depicting the training and testing loss during the training of data done at implementation phase of these methodologies. The followed graph shows the actual speed of wind and predicted one in case of LSTM methodology. At the time of training and testing, certain type of losses occurs, which is due to the uneven occurrence of data. These losses occur due to the glitches in the algorithms.

In the below graphs given in Figure 10 which represents LSTM wind speed forecasting. The dataset is same for both the methodologies so they depict a quite alike graph from each other but different in error rate. These graphs show the uneven vibrations in the speed i.e. in real speed of wind and predicted.

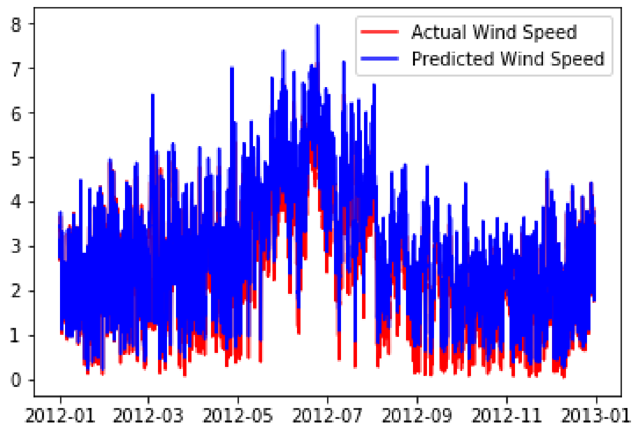

Figure 10. LSTM Wind Forecasting

Figure 11 shows the variations in the actual speed and predicted speed in case of forecasting through SVM algorithms. The vibrations are quite as that of LSTM but the variation in error can be clearly observed while implementation of it on the same dataset.

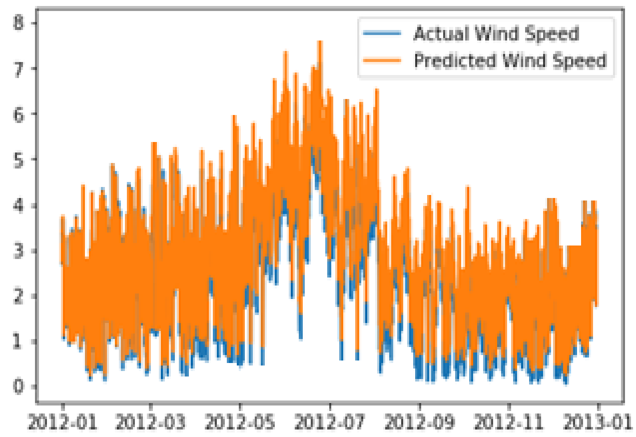

Figure 11. SVM Wind Forecasting

Figure 12 depicts the performance graph for both the techniques (LSTM & SVM). The performance graph is generated by the results which are produced by the both the techniques employed on same data set. The performance parameters for both the algorithms are shown in terms of RMSE value. RMSE for LSTM is 0.427 and RMSE value for SVM is 0.768. It is clearly observed that the error rate is high in SVM as compared to LSTM algorithm. The error rate in SVM is high as it does not hold a property of pattern remembrance for longer durations of time in comparison to LSTM. Both the techniques are quite contrary to each other as SVM works on the theme of hyper planes but LSTM works on the neural network criteria. SVM can be treated as the feed forwarding type network but LSTM can back propagate with time or as required by the system.

So, due to these practices performance rate of LSTM based methodologies are higher than the methodologies comprising of SVM. It is clearly observed by looking the performance graph comprising LSTM to determine speed (wind) is better than the model using SVM technique. Here x-axis shows the error rate and y-axis denotes the algorithm used (LSTM and SVM). The RMSE value is calculated by the formula given in equation 4 and equation 5:

$$MSE_c = \frac{1}{M}\left(\sum_{r=1}^{M} e_{r+c}^2\right) \tag{4}$$

$$RMSE_c = \sqrt{MSE_c} \tag{5}$$

Were e: predicting error
M: total no. of observations

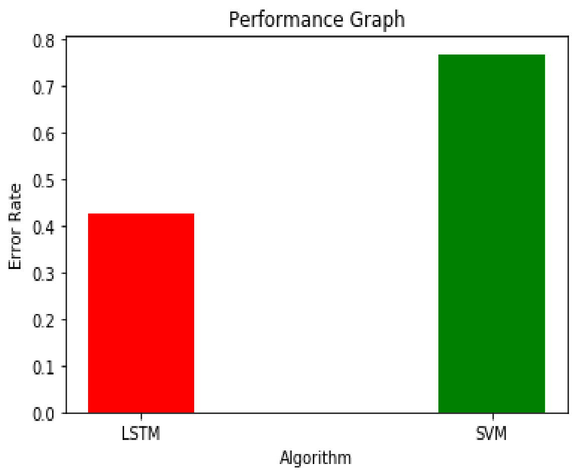

Figure 12. Performance Comparison Graph

5. Conclusion

After the detailed observation of the results from both the methodologies, it can be concluded that the LSTM is more effective as compared to the SVM. The error rate in LSTM is less therefore it can be used more frequently in the forecasting techniques as compared to the other one. LSTM with deep learning can be implemented to obtain more efficient result in the forecasting approaches because of its property of pattern remembrance for longer duration of time. LSTM can be hybridised with other models to generate more accurate models with efficient prediction. After the overall study of both the methodologies the obtained result concludes that LSTM has more significance in forecasting techniques. Thus LSTM with the property of pattern remembrance can be further implemented on larger data sets to obtain the highly accurate results and can be utilised by organisations to predict the better and efficient weather forecasting conditions. In case of wind speed predictions it can be used to maintain the gap between power generation and power utilization.

References

[1] Yesilbudak, M., Sagiroglu, S. and Colak, I. (2013) A new approach to very short term wind speed prediction using k-nearest neighbour classification. *Energy Conversion and Management,* **69**: 77-86.
[2] liu, D., Niu, D., Wang, H. and Fan, L. (2014) Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renewable Energy,* **62**: 592-597.
[3] Pourmousavi Kani, S.A. and Ardehali, M.M. (2011) Very short-term wind speed prediction: A new artificial neural network-Markov chain model. *Energy Conversion and Management,***52**: 738-745.
[4] Jiang, Y., Song, Z., Kusiak, A. (2013) Very short-term wind speed forecasting with Bayesian structural break model. Energy Conversion and Management, **50**: 637-647.
[5] Tronsco, A., Salcedo-sanz, S., Casanova-Mateo, C., Riquelmw, J.C. and Prieto, L. (2015) Local models-based regression trees for very short-term wind speed predictions. *Renewable Energy*, **81**: 589-598.
[6] li, G., shi, J. and Zhou, J. (2011) Bayesian adaptive combination of short –term wind speed forecasts from neural network models. *Renewable Energy*, **36**: 352-359.
[7] liu, H., Mi, X. and Li, Y. (2018) Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolution neural network and convolutional long short term memory network. *Energy conversion and Management*, **166**:120-131.
[8] Cadenas, E. and Rivera, W. (2009) Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. *Renewable Energy*, **34** (1): 274-278.
[9] Cadenas, E. and Rivera, W. (2010) wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. *Renewable Energy*, **35** (12): 2732-2738.
[10] Wang, J., Qin, S., Zhou, Q., China (2015) Medium – term wind speeds forecasting utilizing hybrid models form three different sites in Xinjiang ,China. *Renewable Energy*, **76**: 91-101.
[11] Barbounis, T.G., Theocharis, J.B., Member, IEEE Alexiadis, M.C. and Dokopoulos, P.S., Member, IEEE(2006) Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion*, **21**(1): 273-284.
[12] Kavasseri, R.G. and Seetharaman, K. (2009) Day-ahead wind speed forecasting using f-ARIMA mpdels. *Renewable Energy*, **34** (5): 1388-1393.
[13] Barbounis, T.G. and Theocharis, J.B., (2006) Locally recurrent neural networks for long term wind speed and power predictions. *Neurocomputing*, **69** (4): 466-496.
[14] Gupta, A., Gautam, J. and Kumar, A. (2017) A Survey on the methodologies used for semantic document clustering. *In International Conference on Energy, Communication, Data Analytics and Soft Computing*, Chennai, India, 1-2 Aug. 2017, (IEEE), 671-675.
[15] Kumar, A., Jain, R., Pandey, A., Gupta, S.P. and Garg, S. (2018) Electricity Load Forecasting Methods-A Review. *International Journal of Research in Electronics and Computer Engineering*, **6** (3): 374-378.
[16] Kumar, N., Mishra, B., and Bali, V. (2018) A Novel Approach for Blastit-Induced Fly Rock Predication Based on Particle Swarm Optimization and Artificial Neural Network., *Lecture Notes in Networks and Systems*(Springer Nature Singapore Pte. Ltd), ch. 3.
[17] Jha, P.C. and Bali, V. (2012) Goal Programming Approach for Selection of COTS Components in Designing a Fault Tolerant Modular Software System under Consensus Recovery Block Scheme", *International Journal of Computer and Communication Technology (IJCCT)*, **3** (1): 1-8.
[18] Jha, P.C., Bali, V., Narula, S. and Kalra, M. (2014) Optimal Component Selection Based on Cohesion and Coupling for Component based Software System under Build-or-Buy Scheme. *Journal of Computational Sciences*, **5** (2): 233-242.
[19] Gupta A., Kumar A. and Gautam J. (2017) A survey on semantic similarity measures. *International Journal for Innovative Research in Science & Technology,* **3** (12): 243-247.
[20] Kaur, P., and Sharma, M. (2019) Diagnosis of Human Psychological Disorders using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis. *Journal of Medical Systems,* **43** (7): 204.
[21] Gautam, R., Kaur P., and Sharma M. (2019) A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings. *Progress in Artificial Intelligence*: 1-24.

[22] Sharma, M., Sharma, S. and Singh, G. (2018) Performance Analysis of Statistical and Supervised Learning Techniques in Stock Data Mining. *Data* **3** (4): 54.
[23] Gadekallu, Thippa Reddy, et al. (2019) Application of Data Mining Techniques in Weather Forecasting. *Sentiment Analysis and Knowledge Discovery in Contemporary Business. IGI Global*: 162-174.
[24] Colak, I., Sagiroglu, S., and Yesilbudak, M. (2012) Data mining and wind power prediction: A literature review. *Renewable Energy*, **46**: 241–247.
[25] Han, J., Kamber, M. (2006) Data mining concepts and technique, 1st ed. (San Francisco: Morgan Kaufmann), ch. 3.
[26] Lew, A., Mauch, H. (2006) Introduction to data mining and its applications. *Berlin: Springer:* 1-20