Literature Survey

1. RECURRENT SUPPORT VECTOR MACHINES FOR SPEECH RECOGNITION

   https://www.microsoft.com/en-us/research/wp-content/uploads/2016/06/RecurrentSVM_updated-1.pdf

   They have replaced the softmax layer with SVM classifier at the end of LSTM. LSTM network is used for learning about the data an improving the model, while SVM is used at the last for classifying the data. They have used two methods, frame level training, and sequence level training. The testing is conducted on speech signals. They report minor, but positive improvements in accuracy for LSTM+SVM, compared to LSTM+Softmax, and the basic feed forward neural network.

2. TANDEM LSTM-SVM APPROACH FOR SENTIMENT ANALYSIS

   http://ceur-ws.org/Vol-1749/paper_030.pdf

   They have used similar model as paper no.1, LSTM is used for training and learning the data, and SVM is used at the last for classification. They have used it to classifiy text data (tweets). Bi-directional LSTM has been used to extract features from the data and update the hyper-parameters of the model. SVM then uses these features to classify the data into the correct category. The results show positive improvements with highest accuracy occurring when quadratic SVM is used with LSTM network.

3. AUTOMATIC SLEEP STAGE CLASSIFICATION USING SINGLE CHANNEL ECG

   https://sci-hub.tw/https://doi.org/10.1109/EMBC.2018.8512480

   GRU & LSTM is used together with SVM. Linear and quadratic SVMs both are tested. Bi-directional versions of both the networks are used. ECG data is used to classify sleep patterns. Feature vectors are extracted using RNN (LSTM/GRU) and then passed on to SVM for further classification. Comparison with softmax reveals better accuracy for SVM case. Accuracy for RNN+SVM is greater than that of RNN+softmax. Improvements are most significant & positive (3-4%).

4. LINGUISTIC INFORMED DEEP-LEARNING MODEL FOR IRONY DETECTION

   http://ceur-ws.org/Vol-2263/paper027.pdf

   Multi layered models for LSTM and CNN are made. Text data is used. Softmax layer at last is replaced with linear SVM. Data is pre-processed before entry. LSTM+softmax appears to give better accuracy compared to LSTM+SVM, similar is the case for CNN. The author admits these results are contrary to other findings and suggests better tuning of hyper-parameters in the hidden layers of the network. (Pre-processing and proper tuning are essential) CNN+LSTM+SVM gives better accuracy for SVM compared to softmax.

5. DEEP LEARNING USING LINEAR SUPPORT VECTOR MACHINES

   https://arxiv.org/pdf/1306.0239.pdf

   Proposes similar replacement of softmax classifier with SVM after the final dense layer of the model. Gradient from SVM hinge loss is back propagated to the initial layers for tuning the hyper-parameters. SVM has been used in combination with CNN. Multiclass SVMs have been used (weight regularisation). Categorisation has been done on the MNIST data set. Small, but positive accuracy improvements are recorded (1-2%). CNN+SVM is better than CNN+softmax. Author suggests other ways to implement multiclass SVMs (categorical hinge loss).

6. ECG-BASED BIOMETRICS USING ECURRENT NEURAL NETWORKS

   https://sci-hub.tw/https://doi.org/10.1109/ICASSP.2017.7952519#

   Use of LSTM, GRU and other RNN models for ECG classification is tested. Both, uni-directional and bi-directional networks are analysed. ECG system is used for two different applications, identification and authentication. In identification, given ECG pattern is classified into one of the four categories, whereas in authentication, the ECG pattern is given a class and the system verifies whether it is true or not. Very high accuracy is achieved compared to normal ML networks, however, this is an indicator of over fitting as 100% accuracy is achieved in some cases. Better model is required. LSTM networks are first trained, then validated on some part of the data that had been reserved, and then finally tested on and unseen data set.

7. RNN BASED CLASSIFICATION OF ECG SIGNALS FOR DETECTION OF SLEEP APNEA

   https://www.researchgate.net/profile/Adil_Khan13/publication/319056221_Recurrent_Neural_Network_Based_Classification_of_ECG_Signal_Features_for_Obstruction_of_Sleep_Apnea_Detection/links/59e0209faca272386b634713/Recurrent-Neural-Network-Based-Classification-of-ECG-Signal-Features-for-Obstruction-of-Sleep-Apnea-Detection.pdf

LSTM Network is used for detecting apnea from an ECG signal. Time interval between two R R waves is used to get the signal. The features of the data are extracted using LSTM model with 3 hidden layers and one final output layer. SVM and other classifiers are tested at the output layer. Best accuracy is achieved for LSTM+SVM, but use of 3 layers and improper regularisation (linear instead of quadratic), leads to overfitting, or underfitting. Better model can be made with improvements in these areas.

8. ANOMALY DETECTION IN ECG SIGNAL USING LSTM NETWORKS

   https://www.researchgate.net/profile/Sucheta_Chauhan3/publication/308852664_Anomaly_detection_in_ECG_time_signals_via_deep_long_short-term_memory_networks/links/5a719365458515015e64a698/Anomaly-detection-in-ECG-time-signals-via-deep-long-short-term-memory-networks.pdf

   Single channel ECG data is used for detection of anomalies in ECG Data. Deep LSTM network by stacking multiple LSTM layers is used. Softmax classification at output layer is used. Accuracy achieved is good, but use of deep lstm networks greatly increases computational time, and leads to a higher load. Usage of deep lstm networks is required in a more efficient manner.

9. RNN REPLACEMENT OF SOFTMAX WITH SVM

   https://arxiv.org/pdf/1709.03082.pdf

   GRU is combined with svm for intrusion detection in network traffic data. GRU network is used, wherein the GRU layer is used for feature extraction, and svm is used for the final classification. Comparison is made between GRU with softmax output layer and GRU with svm output layer. Higher accuracy is achieved for the GRU+SVM case by about 3-4 %. The training time was also less for the svm which shows that it is computationally less intensive. However, use of GRU is not desirable, as even higher accuracy can be achieved by using LSTM network along with SVM classifier.

10. DEEP-LEARNING BASED STRESS DETECTION ALGORITHM WITH SPEECH SIGNAL

    https://sci-hub.tw/10.1145/3264869.3264875

    Stress is detected in speech signals by first extracting features of the data using LSTM and other feed forward networks, and then classifying them into two categories, stressed or unstressed using different classifiers in the output layer of the network. Softmax and SVM classifiers are tested in the final layer. Four different stress detection modules are used for analysis. There are 40 input nodes, and 2 possible outputs. SVM classifier gives higher accuracy by about 2-3%. Sentence Level features perform. However, better regularisation techniques can be used, apart from using just dropout. Model can be modified further to extend it to multiclass classification, as is the case for ECG and other health data.

11. DETECTION OF HYPOVIGILANCE USING POSTURE CONTROL SYSTEM MEASURES

    https://s3.amazonaws.com/academia.edu.documents/7821110/MB2010_Proceedings_web.pdf?response-content-disposition=inline%3B%20filename%3DBall_Recovery_in_the_Handball_Tournament.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F20190719%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20190719T070800Z&X-Amz-Expires=3600&X-Amz-SignedHeaders=host&X-Amz-Signature=4a923d557b04188e6e3da2676318f6396ba0196d0544ab0cfd3af5cc4d362702#page=201

    Posturographical data is used to predict alertness state of a driver, and classify it into two categories, vigilant, and hypovigilant. Two feature sets are extracted, on in time domain, and the other in spectral domain. Feature extraction and learning is done by various kinds of neural networks. Final classification is binary and is done by output layer. SVM is used standalone, and combined with another networks too. Error percentage is calculated, and it is the lowest for LSTM+SVM compared to LSTM+softmax, and other ANNs. This is however, only applicable to time domain data. Still, better ways can be used to optimise the LSTM network parameters, as well as the weight values. Author suggests that this method is best used only for short data sets, and better models are required for large posturographical data sets.

12. LSTM BASED AUTO-ENCODER MODEL FOR ECG ARRHYTHMIAS CLASSIFICATION

    Combination of lstm network and svm classifier is used for arrhythmia detection. LSTM based auto encoder learns and extracts the features from the data, and SVM is used to classify the signals from the learned features into 5 categories. MIT-BIH dataset is used. RR interval between two heartbeats is one of the features that have been used. Extraction methods such as principal component analysis, wavelet transform have been applied. Wavelet method is primarily used. Two LSTM networks are used, where one works as the encoder, and the other as the decoder. Original and denoised data both are tested. Highest accuracy is achieved when LSTM+SVM is used. K fold cross validation and denoised data give better results. However, the data across the heartbeats is imbalanced, and this leads to nearly 100% accuracy, which indicates overfitting. The model needs to be updated with dropout and other such techniques to prevent this.

13. DECISION MAKING METHOD FOR DIAGNOSIS OF THYROID BASED ON DEEP LEARNING

    https://sci-hub.tw/10.1145/3231884.3231893

    Deep learning models based on LSTM networks are used to assess the severity of thyroid and give advice according to which the level of treatment and hospital care in the hierarchical system is prescribed. Data about the patient is taken from blood test and ultrasound test. This data is then passed to the LSTM network (or CNN) which processes it and provides the feature vector. That is then passed to the last layer, either

softmax, or SVM classifier for final classification and recommendation of healthcare. Highest accuracy is achieved for LSTM+SVM case, greater than LSTM+softmax or the CNN variants. During processing though, data preprocessing has not been done beforehand, this leads to lower than expected results. More number of nodes, or a better defined LSTM network with dropout and other regularisation features might give better results.

14. HYBRID ATTENTION BASED MULTIMODAL NETWORK FOR SPOKEN LANGUAGE CLASSIFICATION

    https://www.aclweb.org/anthology/C18-1201

    LSTM Networks are used for classifying vocal signals/ spoken language, specifically medical speech from trauma situations. Comprehension of speech signal is difficult due to the multiple heterogenous inputs. Data is taken from various audio files, and text files. Data preprocessing is done using word2vec.Modality attention fusion is used. It is then passed to various kinds of networks (LSTM,CNN,etc) for feature extraction. Data is divided into test and train datasets. A combination of LSTM and SVM provides good accuracy compared to lstm+softmax, but other networks are able to give even higher accuracy. This suggests that better high level feature extraction needs to be done, and regularisation features, dropout, etc need to be used to get better results.

15. TOPIC CLASSIFICATION BASED ON DISTRIBUTED DOCUMENT REPRESENTATION AND LATENT TOPIC INFORMATION

    http://www.apsipa.org/proceedings/2017/CONTENTS/papers2017/13DecWednesday/Poster%202/WP-P2.12.pdf

    Natural language processing is done to assign topics to documents. Topic modelling method of Latent Dirichlet Allocation (LDA) is used. In this, the documents are represented as low dimensional vectors. CNN and other RNN based models are tested. The model (CNN/RNN) is trained on the data and then classification is done by the final layer. In this case, accuracy for LSTM+SVM is lower than LSTM + softmax, although it is higher than CNN and its variants. Highest accuracy is achieved for the LSTM+LDA model. However, it is computationally more intensive and even higher accuracy can be achieved for LSTM variants through better use of regularisation and LSTM specific data pre processing.