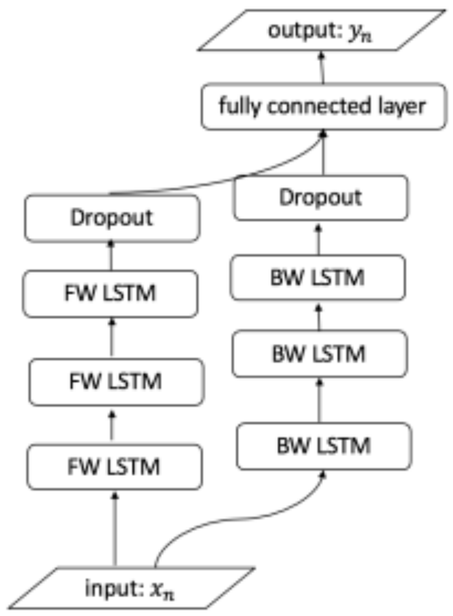
Ensemble Model-Based Prediction Using Recurrent Neural Networks for Electronic Medical Record

Donghoon Oh, Seungtae Kang, Sangwon Lee and Gil-Jin Jang
School of Electronics Engineering, Kyungpook National University, Daegu 41566, Korea

Introduction

- **EMR (Electronic Medical Record) is a computerized chart of a patient, which includes various personal information, examination and treatment results, operation records and medical history.**
- **RNN (Recurrent Neural Network) model is used to predict disease inferred by the patient's previous history and disease.**
- **Ensemble of many different neural networks generate more accurate results by fusing each learned model.**

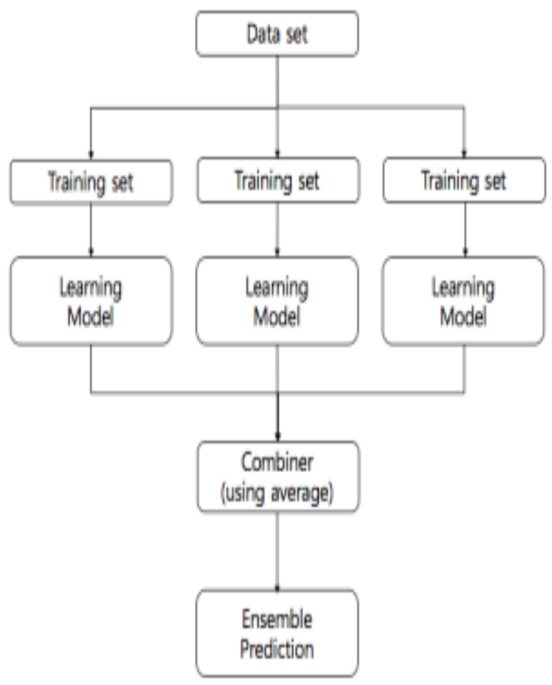
RNN model

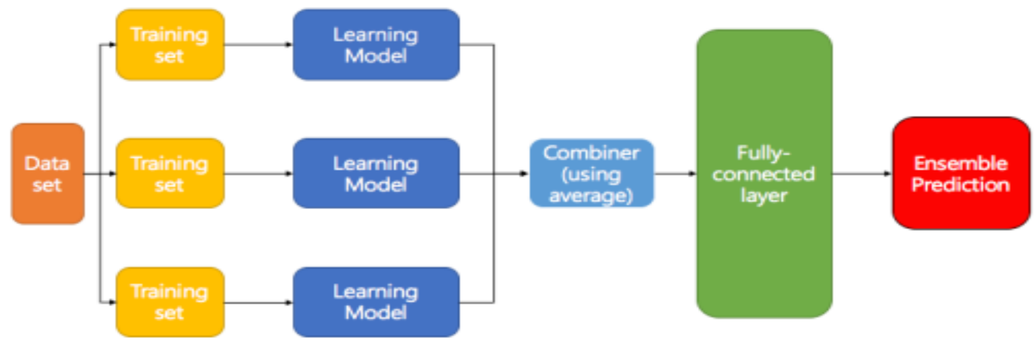


- Bi-directional LSTM (BLSTM)
- Add dropout to last layer
- Combine output of BLSTM
- Fully-connected to the final output

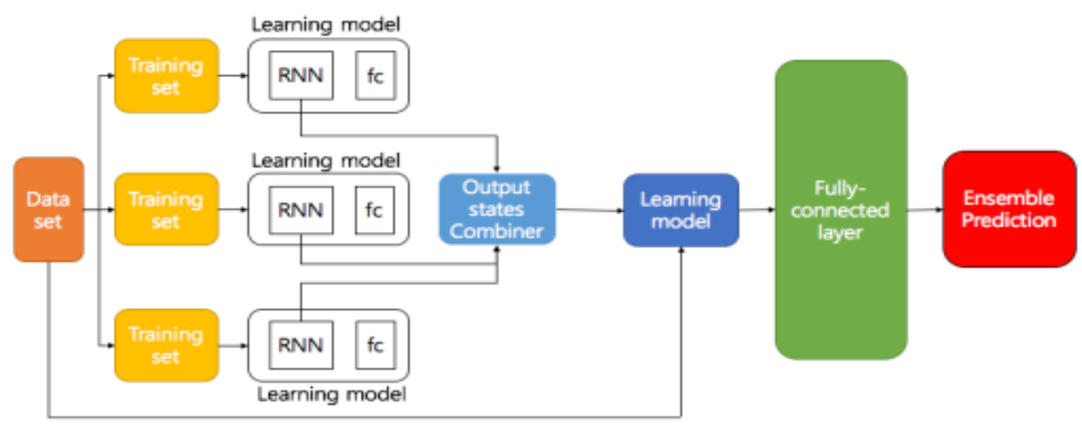
ENSEMBLE MODEL

Bagging





Bagging + fully-connected layer



RNN state output averaging + fully-connected layer

Electronic Medical Record

Patient information from 2002 to 2013
-> Predict data for 2013 using data from 2002 to 2012

Total patients	Train patients	Test patients
11692	8184	3508

list	abbreviation
Systolic blood pressure	BP_HIGH
Diastolic blood pressure	BP_LWST
Blood sugar	BLDS
Total cholesterol	TOT_CHOLE
hemoglobin	HMG
Urine protein	OLIG_PROTE_CD*2
(Serum GOT) AST	SGOT_AST
(Serum GPT) ALT	SGPT_ALT
Gamma GTP	GAMMA_GTP

*2 Since urine protein is a value expressed as an integer. It is not suitable for the inference method proposed in this study.

Not prediction values just used learning

list	abbreviation
Height	HEIGHT
Weight	WEIGHT
Sex	SEX
Age	AGE_GROUP

Experiments Result

Normalized RMSE (Root Mean Square Error) Comparison for Normal and Ensemble Models

LSTM

BP_HIGH	BP_LWST	BLDS	TOT_CHOLE	HMG	SGOT_AST	SGPT_ALT	GAMMA_GTP
0.117	0.127	0.237	0.203	0.121	0.422	0.535	0.767

BLSTM

BP_HIGH	BP_LWST	BLDS	TOT_CHOLE	HMG	SGOT_AST	SGPT_ALT	GAMMA_GTP
0.112	0.124	0.235	0.177	0.115	0.421	0.534	0.766

Ensemble (bagging)

BP_HIGH	BP_LWST	BLDS	TOT_CHOLE	HMG	SGOT_AST	SGPT_ALT	GAMMA_GTP
0.112	0.123	0.235	0.176	0.107	0.421	0.534	0.767

Ensemble (bagging + FC layer)

BP_HIGH	BP_LWST	BLDS	TOT_CHOLE	HMG	SGOT_AST	SGPT_ALT	GAMMA_GTP
0.111	0.122	0.234	0.177	0.108	0.421	0.534	0.767

Ensemble (RNN state output averaging + FC layer)

BP_HIGH	BP_LWST	BLDS	TOT_CHOLE	HMG	SGOT_AST	SGPT_ALT	GAMMA_GTP
0.109	0.120	0.233	0.175	0.096	0.420	0.533	0.767

Conclusion

- ❑**We confirmed the performance of EMR data prediction using RNN.**
- ❑**Ensemble of neural networks performs better than each RNN models to prediction.**
- ❑**Furthermore, continuous improvement of performance can be expected by developing an algorithm that utilizes newly added EMR information for re-learning (adaptation) of already learned models.**