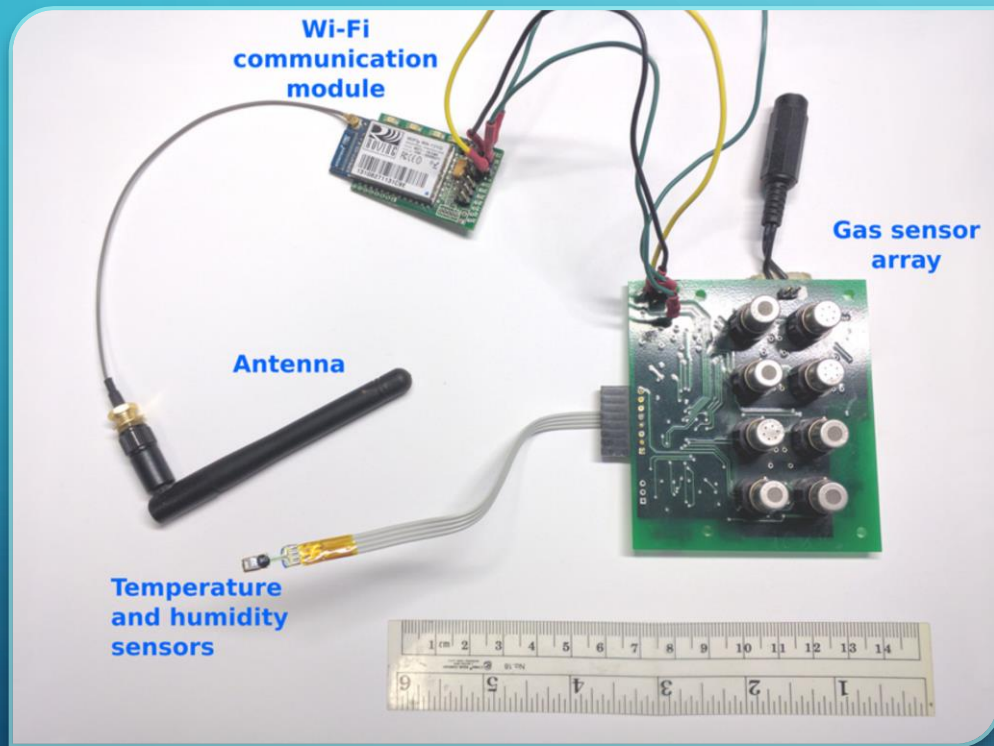


A decorative graphic on the left side of the slide, consisting of a network of white lines and small circles on a blue gradient background, resembling a circuit board or neural network structure.

EE660 FINAL PROJECT

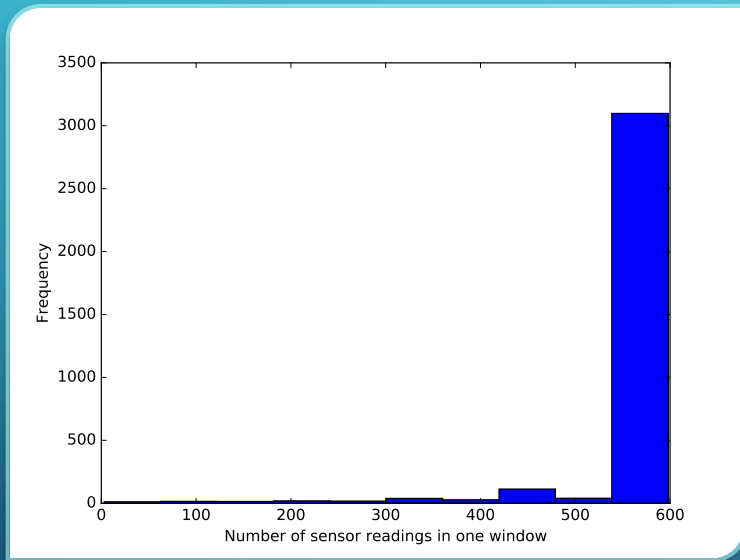
CLASSIFY TIME SERIES DATA WITH RECURRENT NEURAL NETWORK

PROBLEM DESCRIPTION

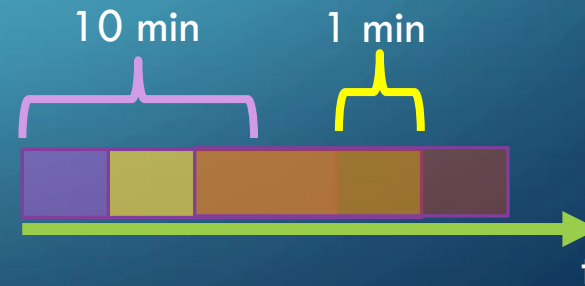


- Time series data from gas sensor array
- Three different presentations: banana, wine and background
- Method used in paper: iSVM

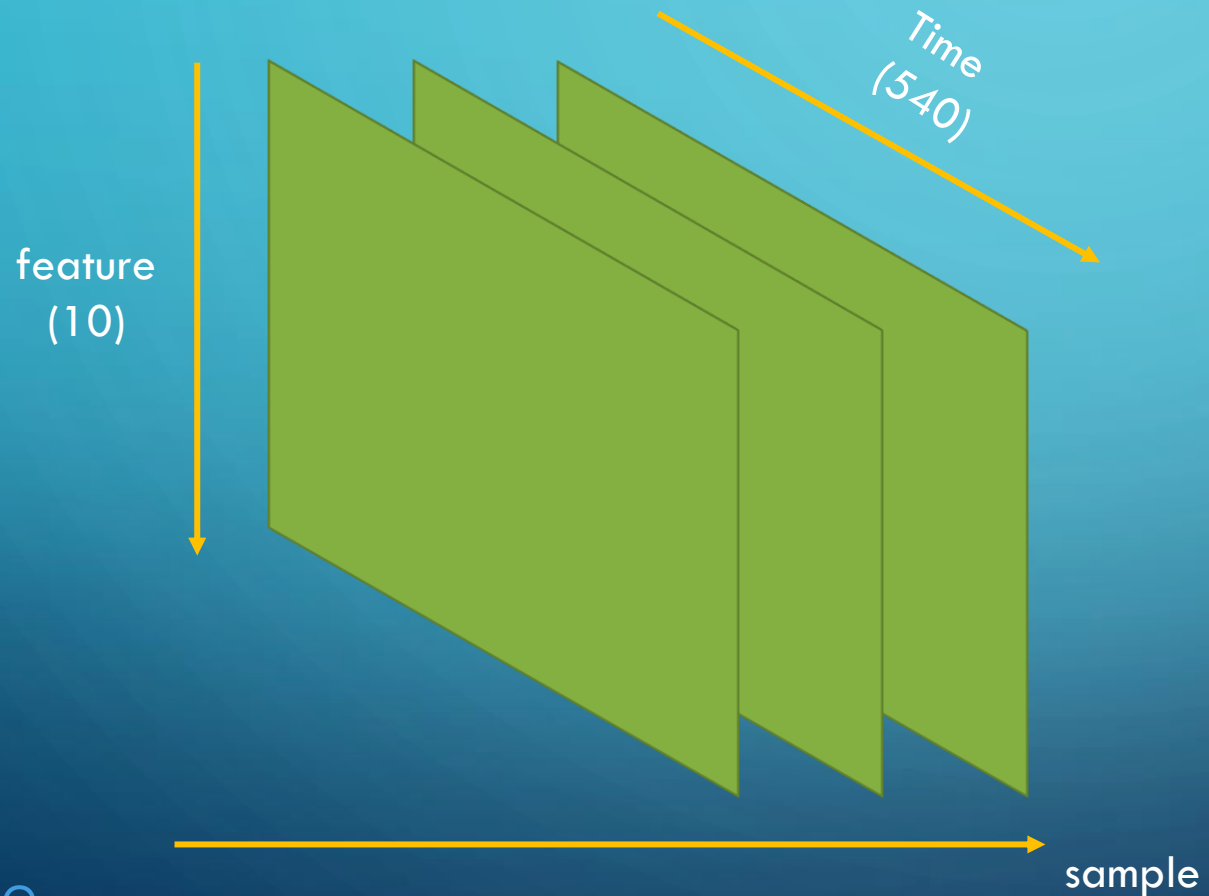
DATA PREPROCESSING



- 99 presentation
- Discard window without enough data point
- Truncate window length to 540



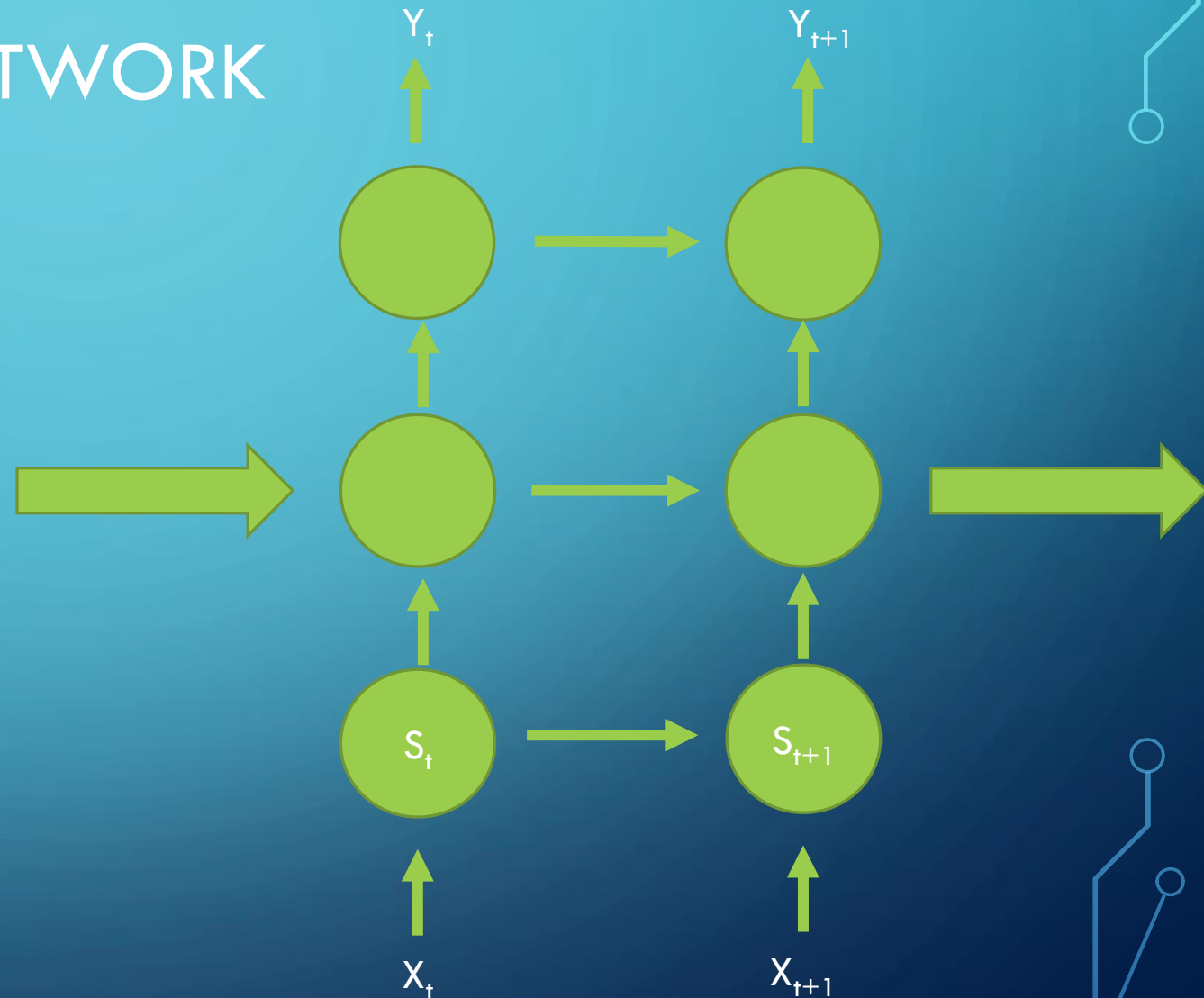
DATA DIMENSION



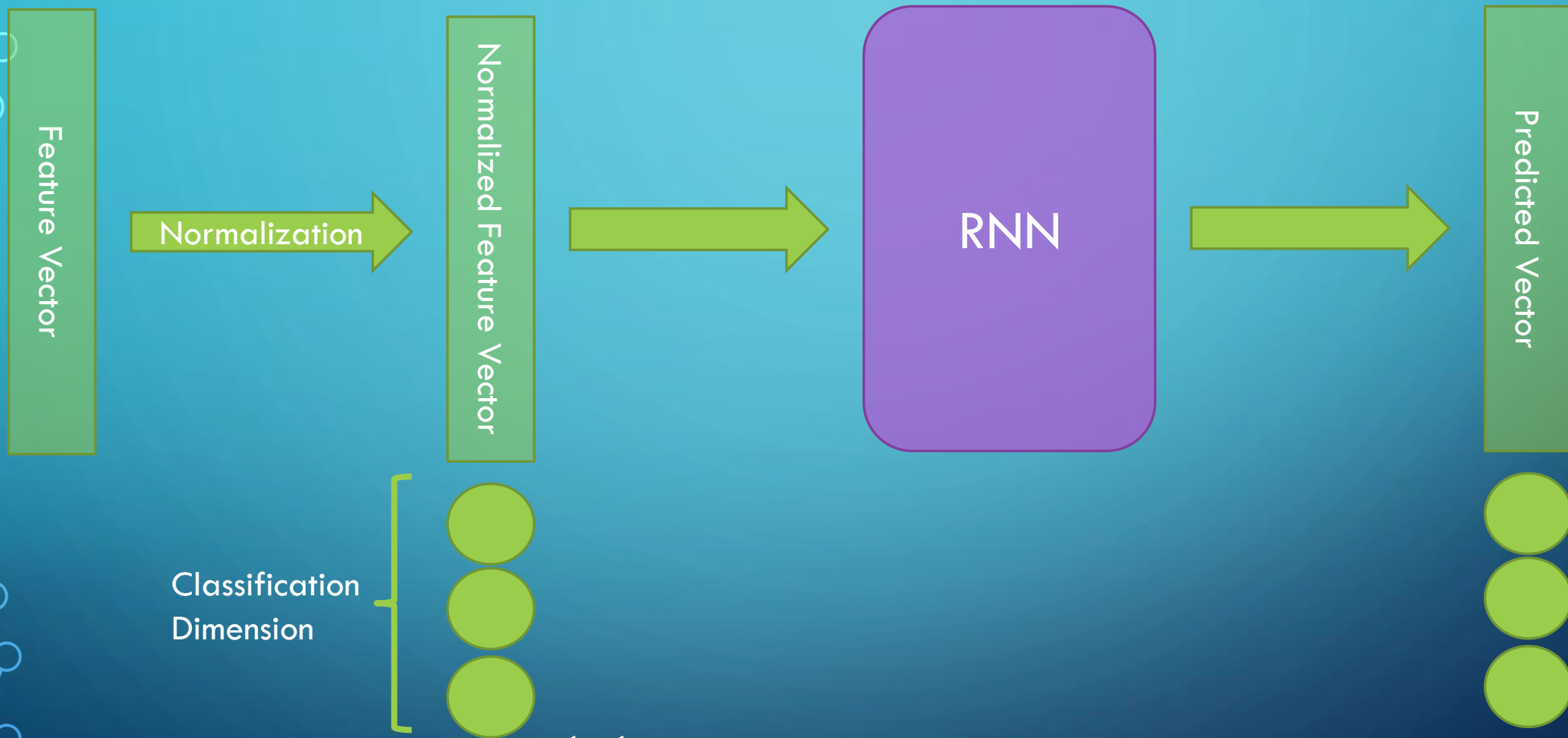
- With a total of 3098 samples

RECURRENT NEURAL NETWORK

- Good at predicting time series
- Long Short Term Memory Architecture (LSTM)

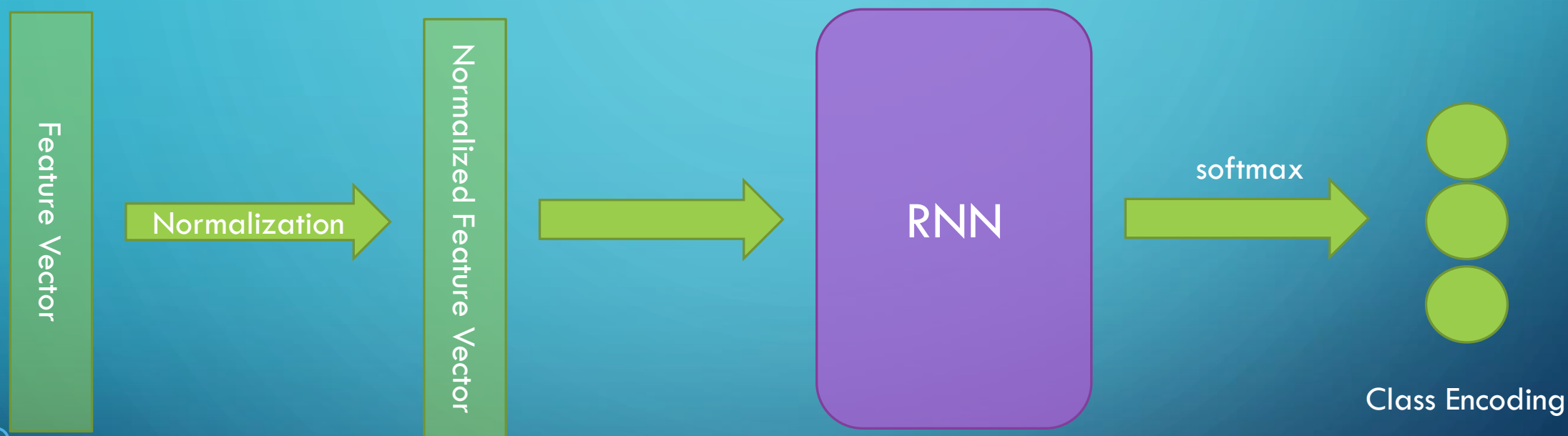


ONE FOR ALL ENCODING



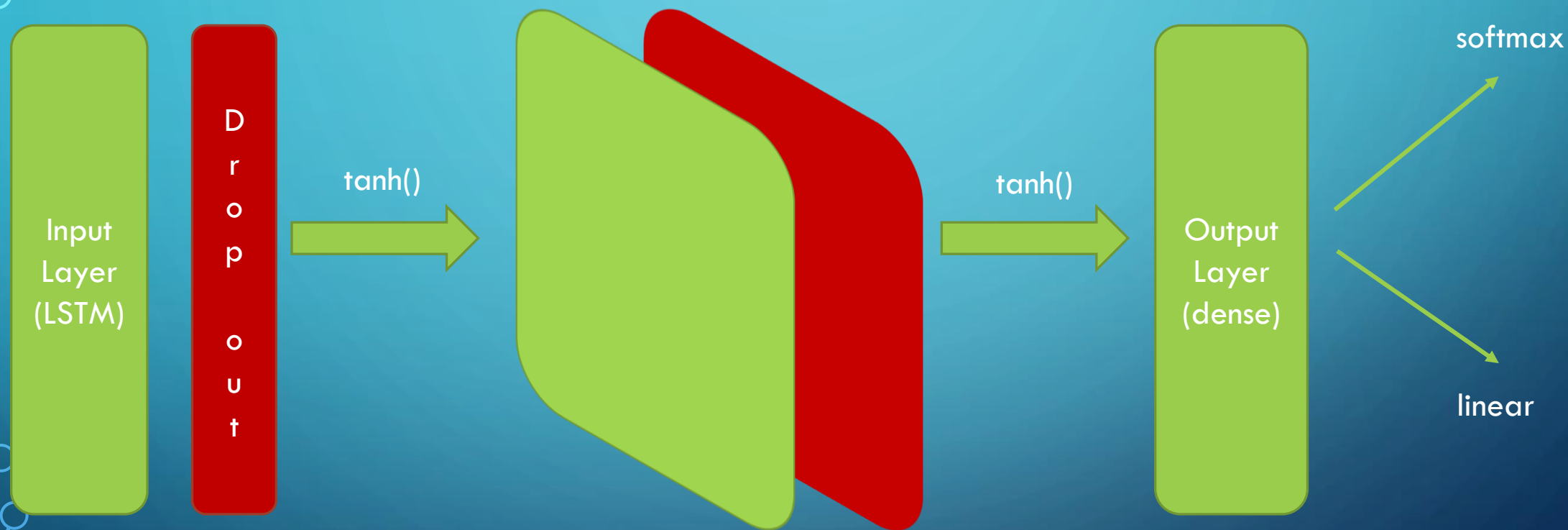
$$L(\vec{w}) = \frac{1}{N} \frac{1}{T} \sum_N \sum_T (\hat{y}_c - y_c)^2 + (\hat{y}_p - y_p)^2$$

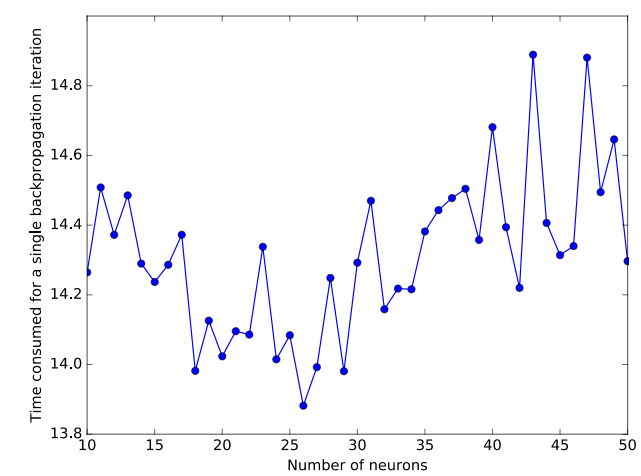
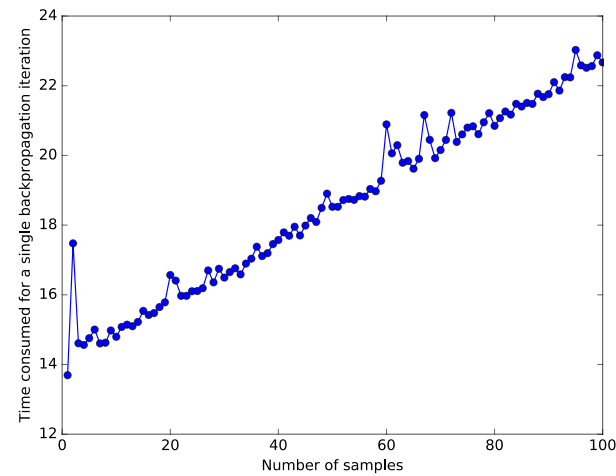
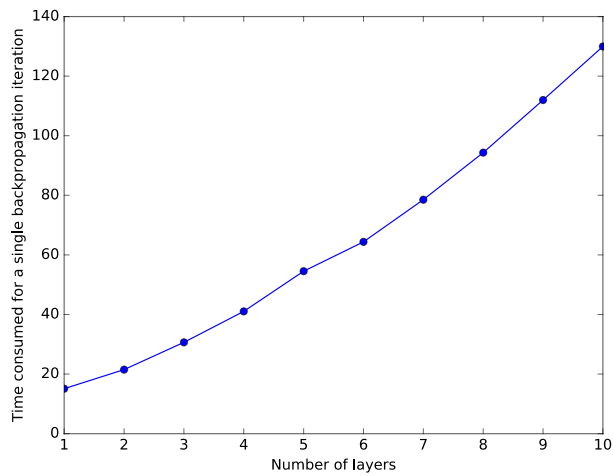
STANDARD CLASSIFICATION



$$L(\vec{w}) = \frac{1}{N} \frac{1}{T} \sum_N \sum_T H(\hat{y}_c, y_c)$$

RNN ARCHITECTURE





COMPLEXITY ANALYSIS

TWO SETTINGS

- Randomly split the 99 presentations into training 79 training set and 20 testing set. (0.8 split ratio) When training, feed the time windows in sequence into RNN.
- Randomly split 3098 windows into 2478 training set and 620 testing set. When training, shuffle all the windows and then feed them into RNN.

RESULTS

Split with respect to groups; One for all encoding

	Number of Hidden Layers	Neurons per Layer	Mean Misclassification Rate	Standard Deviation
1	1	10	0	0
2	2	10	0.002	0.006
3	1	14	0	0
4	2	14	0	0

The results are averaging over 10 different training and testing split

RESULTS

Split with respect to groups; Classification only

	Number of Hidden Layers	Neurons per Layer	Mean Misclassification Rate	Standard Deviation
1	1	10	0.0029	0.0077
2	1	15	0.0012	0.0033
3	1	20	0.0054	0.0021
4	2	10	0.018	0.056
5	2	11	0.0023	0.0074
6	2	12	0.023	0.066

The results are averaging over 10 different training and testing split

RESULTS

Split with respect to windows; One for all encoding

	Number of Hidden Layers	Neurons per Layer	Mean Misclassification Rate	Standard Deviation
1	1	10	0.27	0.033
2	1	13	0.22	0.04
3	1	14	0.19	0.032
4	2	10	0.23	0.05
5	2	13	0.18	0.026
6	2	14	0.185	0.05

The results are averaging over 20 different training and testing split

COMPARISON WITH EXISTING RESULT

Feature set	Cross-validated accuracy	Accuracy in test	Std	<i>p</i> -Value
RS	78.5%	76.5%	6.8%	0.02*
RS,T,H	73.3%	71.1%	6.8%	$1 \cdot 10^{-12}$ **
FS	72.4%	71.2%	4.8%	$2 \cdot 10^{-12}$ **
RS,FS	82.6%	80.9%	6.3%	1

FINAL THOUGHT

