

QUESTION ANSWERING BASED SYSTEM

Submitted By:

Chanchal Suman(1611CS01)

Himani Srivastava(1611CS04)

PROBLEM STATEMENT

- ◉ We have implemented an Question Answering based System, It's a Closed domain system Question asked from the given dataset only.
- ◉ We have used bAbi dataset for our project which is used by facebook .

LITERATURE SURVEY

1. Deep Learning for Question Answering By Mohit Iyyer
2. Question Answering using Deep Learning By Eylon Stroh and Priyank Mathur.
3. Jason Weston, Antoine Bordes, Sumit Chopra, Tomas Mikolov, Alexander M. Rush, "Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks"
4. Jason Weston, Antoine Bordes, Sumit Chopra, Tomas Mikolov, Alexander M. Rush, "Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks".

PAPER 1	RECURRENT NEURAL NETWORK	95%
PAPER 2	RECURRENT NEURAL NETWORK(GRU)	99%
PAPER 3	Co-attention neural network	82%(UNDER PROGRESS)
PAPER 4	RECURRENT NEURAL NETWORK(LSTM)	98.6

DATASET

- ◉ We have used bAbi dataset which is developed by facebook that contain different single and multiple supported fact based dataset , We have implemented our project based on single supported fact only.

- ◉ Babi DataSet Sample:

1 Mary moved to the bathroom.

2 John went to the hallway.

3 Where is Mary? Bathroom 1

4 Daniel went back to the hallway.

5 Sandra moved to the garden.

6 Where is Daniel? hallway 4

IMPLEMENTATION

- ◉ We have trained our network on 10000 trained data set .
- ◉ Test is being done on 1000test data set.
- ◉ We have used RNN(LSTM) .
- ◉ We have used Sigmoid and RELU as activation function.
- ◉ We have used adam and rmsprop as an optimizer.

IMPLEMENTATION...

◉ Tokenize()

('Jaya went to the temple. Where does Jaya gone?')

['Jaya' , 'went' , 'to' , 'the' , 'temple' , '.' , 'Where' , 'does' , 'Jaya' , 'gone' , '?']

◉ Pruning()

>>>>line: Where is John? bedroom 8

Question:Where is John?

Answer: bedroom

Supporting: 8

IMPLEMENTATION

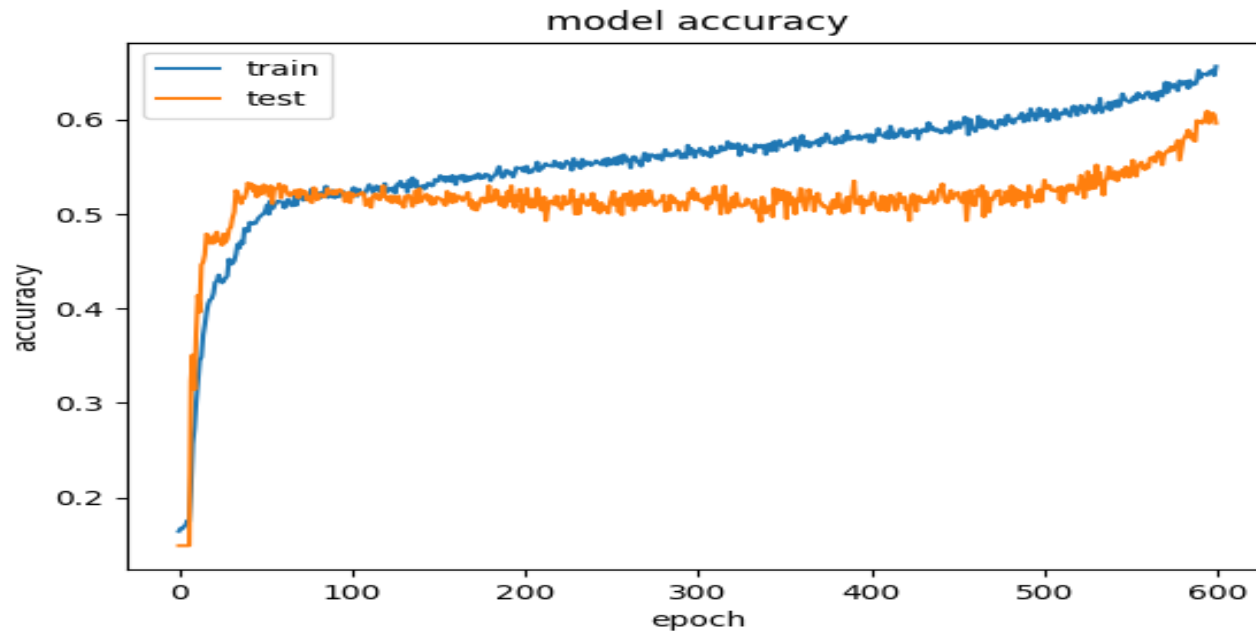
- ◉ **Vectorize:**[12,9,11,14.....]
- ◉ **QuestionArray::** Where is john ?>>>[7 13 4 2]
- ◉ **AnswerArray::**Bathroom >>[0 0 0 0 0 0 0 0 1 0....]

VocabularyId:

{u'hallway': 12, u'bathroom': 9, u'garden': 11, u'journeyed': 14, u'office': 17, u'is': 13, u'bedroom': 10, u'moved': 16, u'back': 8, u'went': 21, u'.': 1, u'to': 19, u'travelled': 20, u'Daniel': 3, u'Sandra': 6, u'the': 18, u'John': 4, u'Where': 7, u'Mary': 5, u'?: 2, u'kitchen': 15}

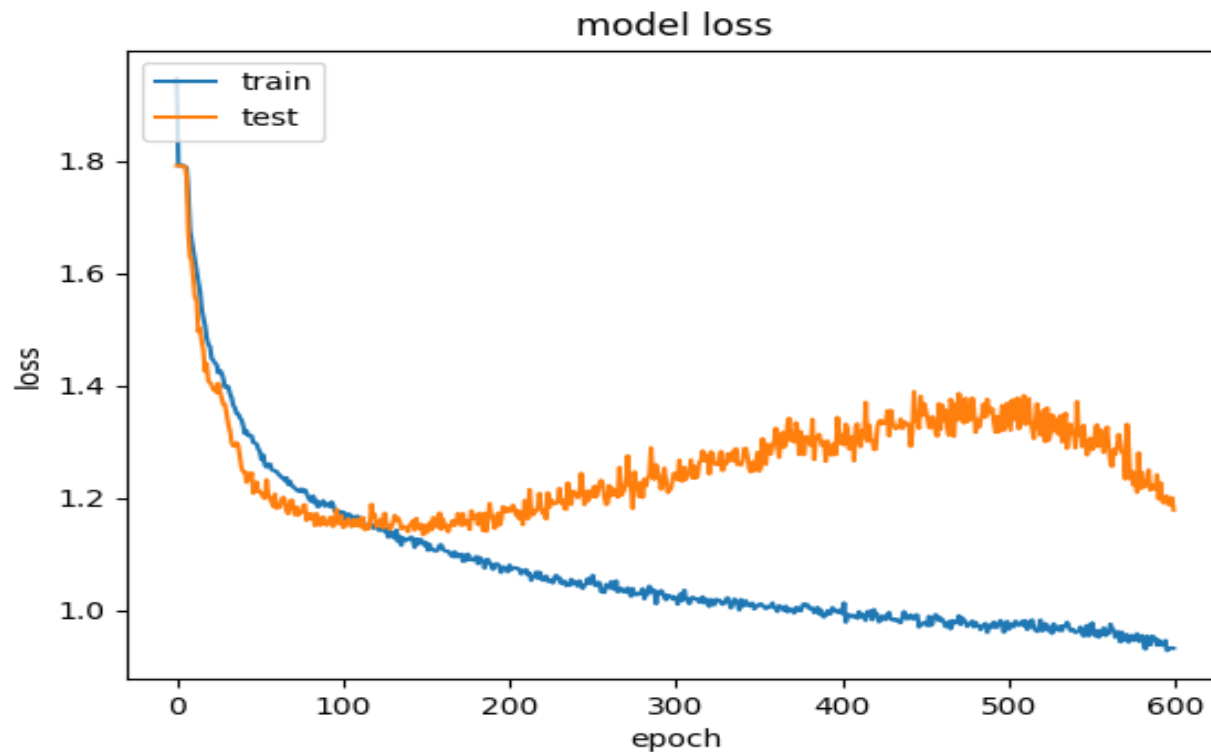
OUTPUTS

- AccuracyVsEpoch(Sigmoid, Epoch600, Dropout =0.7)



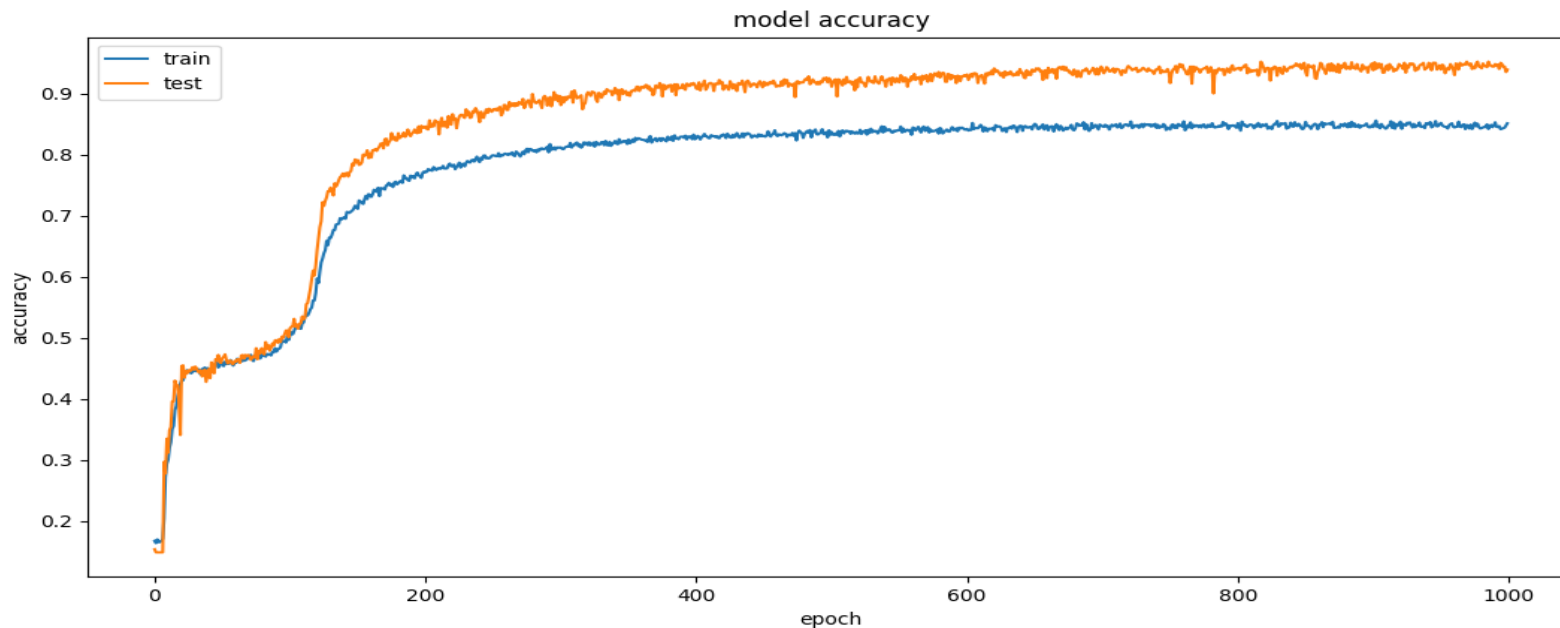
OUTPUTS

- LossVsEpoch(Sigmoid, Epoch600, Dropout=0.7)



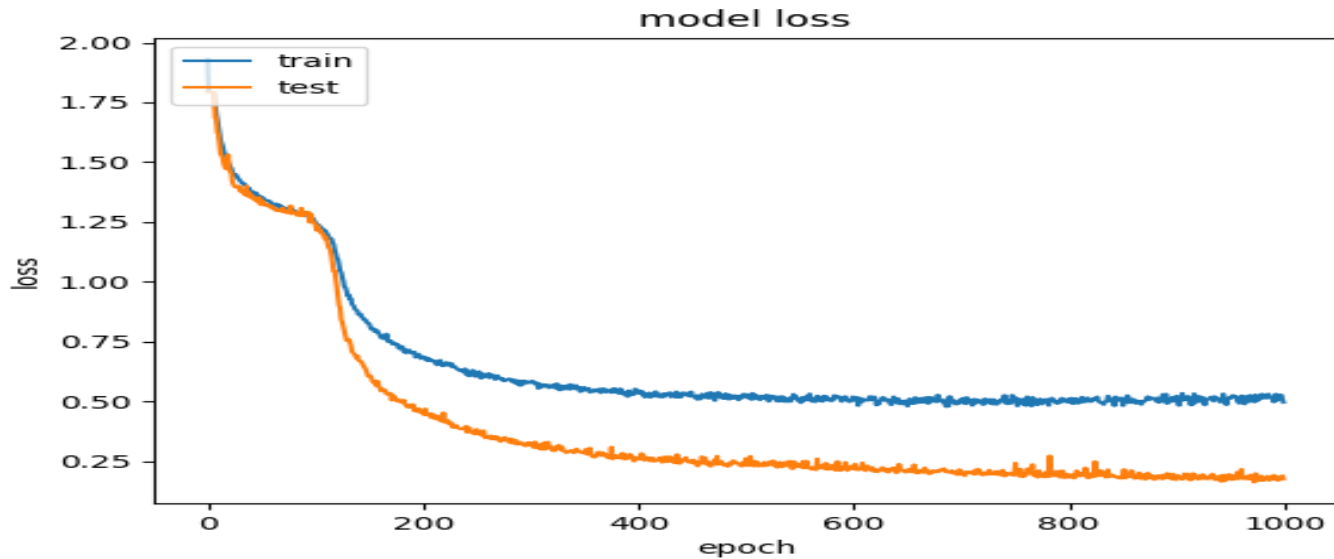
OUTPUTS

- Accuracy Vs Epoch
(Sigmoid, Dropout=0.7, Epoch1000)



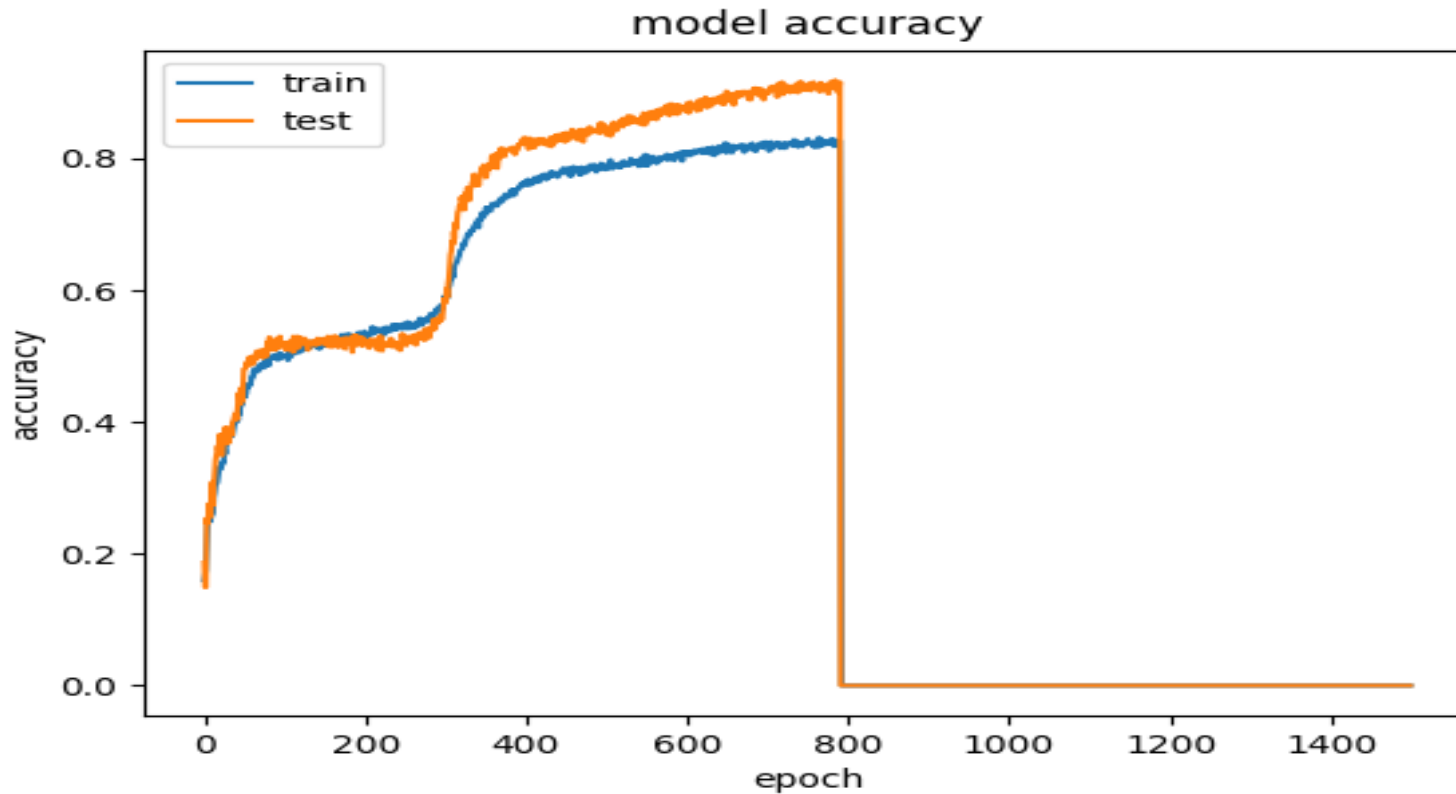
OUTPUTS

- LossVsEpoch(Sigmoid, EPoch1000, Dropout0.7)

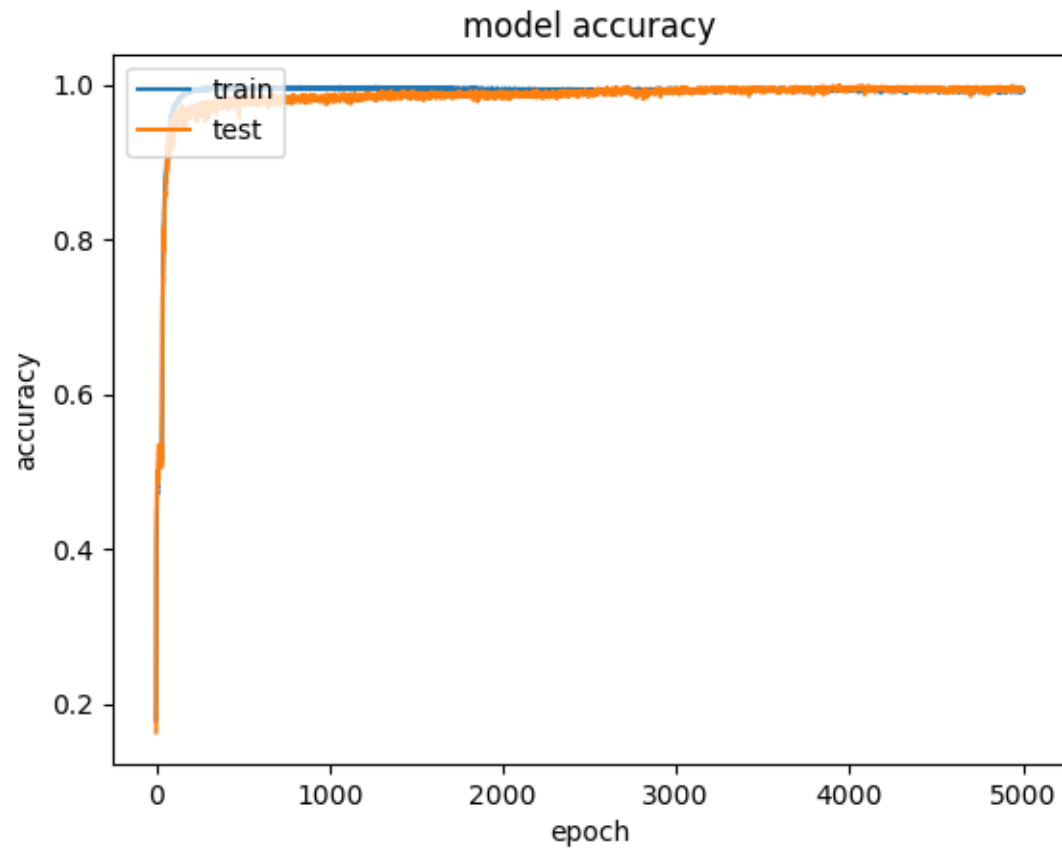


OVERFITTING

AccuracyVsEpoch(Sigmoid, Epoch1500)



RELU WITHOUT OVERFITTING...



OBSERVATIONS

ACTIVATION FUNCTION	OPTIMIZER	EPOCH	DROPOUT	ACCURACY
RELU	RMS-PROP	600	0.3	98.70
RELU	RMS-PROP	1000	0.3	99.5
RELU	RMS-PROP	1500	0.3	99.78
SIGMOID	RMS-PROP	600	0.7	59.6
SIGMOID	RMS-PROP	1000	0.7	93.9
SIGMOID	RMS-PROP	1500	0.7	OVERFITTING
SIGMOID	ADAM	1000	0.7	95.10
RELU	RMS-PROP	5000	0.3	99.6

CONCLUSION

- We have simulated above results by changing activation function to RELU, dropout as 0.3, Optimizer as adam optimizer and observed that using RELU we got better and fast accuracy than Sigmoid but till 5000 epoch we checked and wouldn't be able to get overfitting case, also using Adam Optimizer gives a very bad accuracy.

REFERENCES

- ◉ Jason Weston, Antoine Bordes, Sumit Chopra, Tomas Mikolov, Alexander M. Rush, **"Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks"**,
- ◉ Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, Rob Fergus, **"End-To-End Memory Networks"**,

THANKYOU

◉ QUERIES????