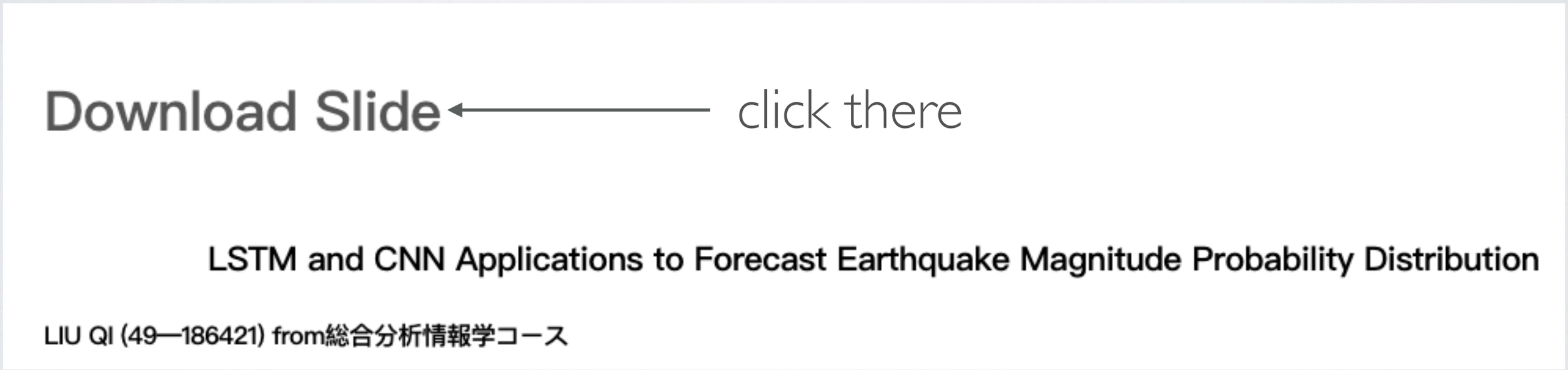
Download slide:

"https://liuqi6776.github.io/2019/04/02/research-expect/"



LSTM AND CNN APPLICATIONS TO FORECAST EARTHQUAKE MAGNITUDE PROBABILITY DISTRIBUTION

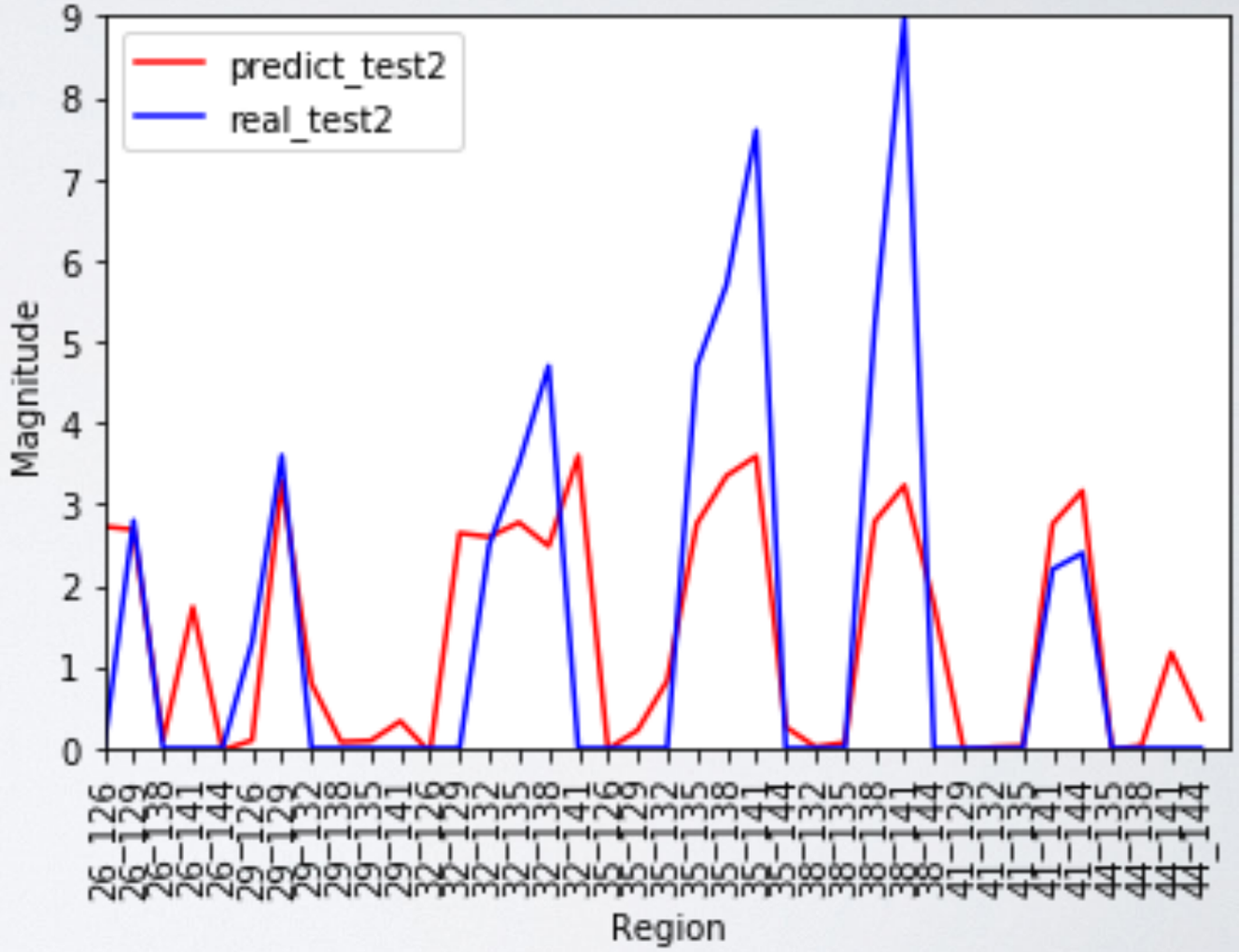
QI LIU
GSII :Applied Computer Science Course
ID:49186421

CONTENTS

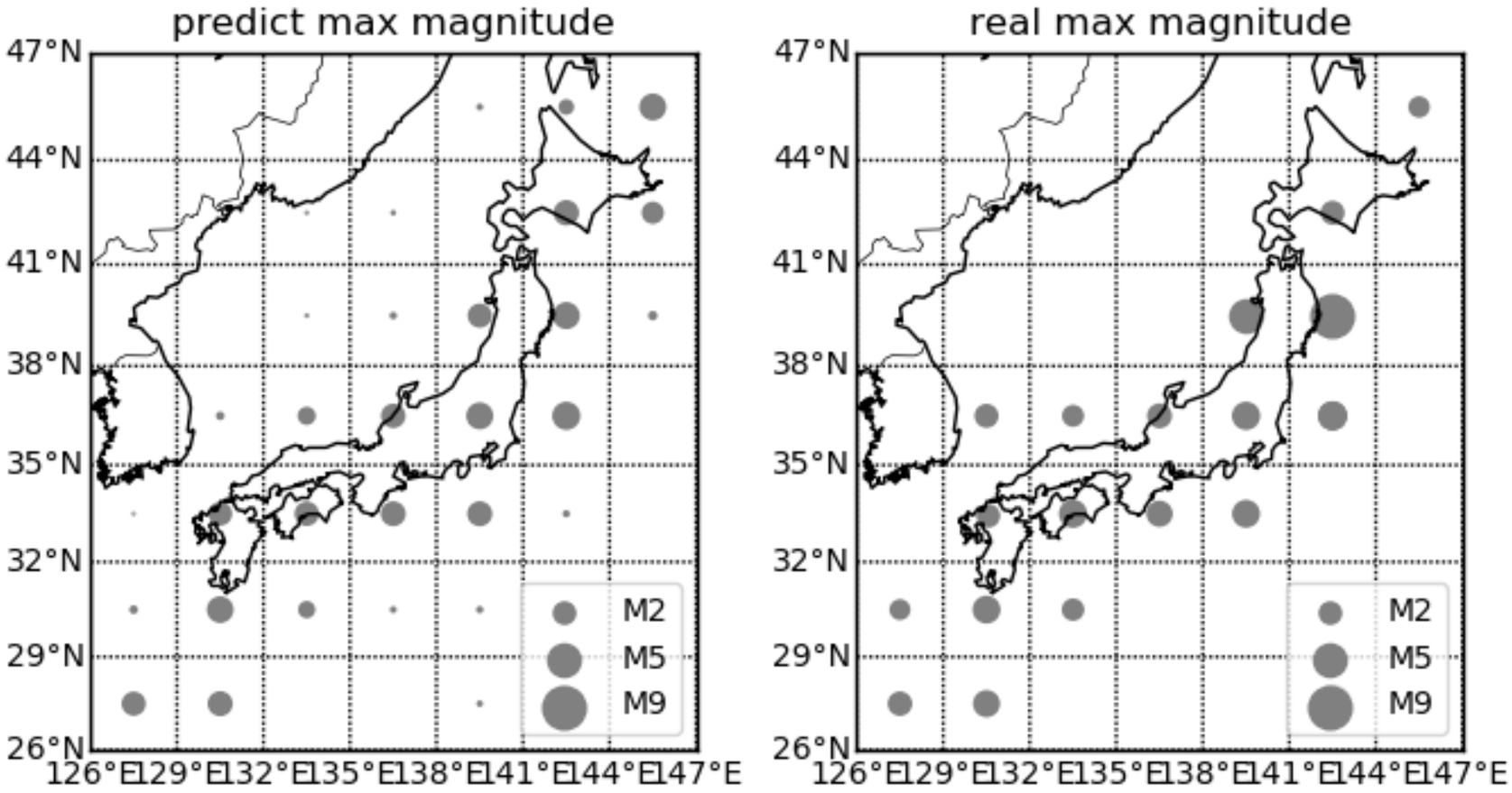
BRIEF INTRODUCTION

About earthquake prediction:

The Tohoku earthquake
Three points:
1,WHEN
2,WHERE
3,WHAT(magnitude)



BRIEF INTRODUCTION

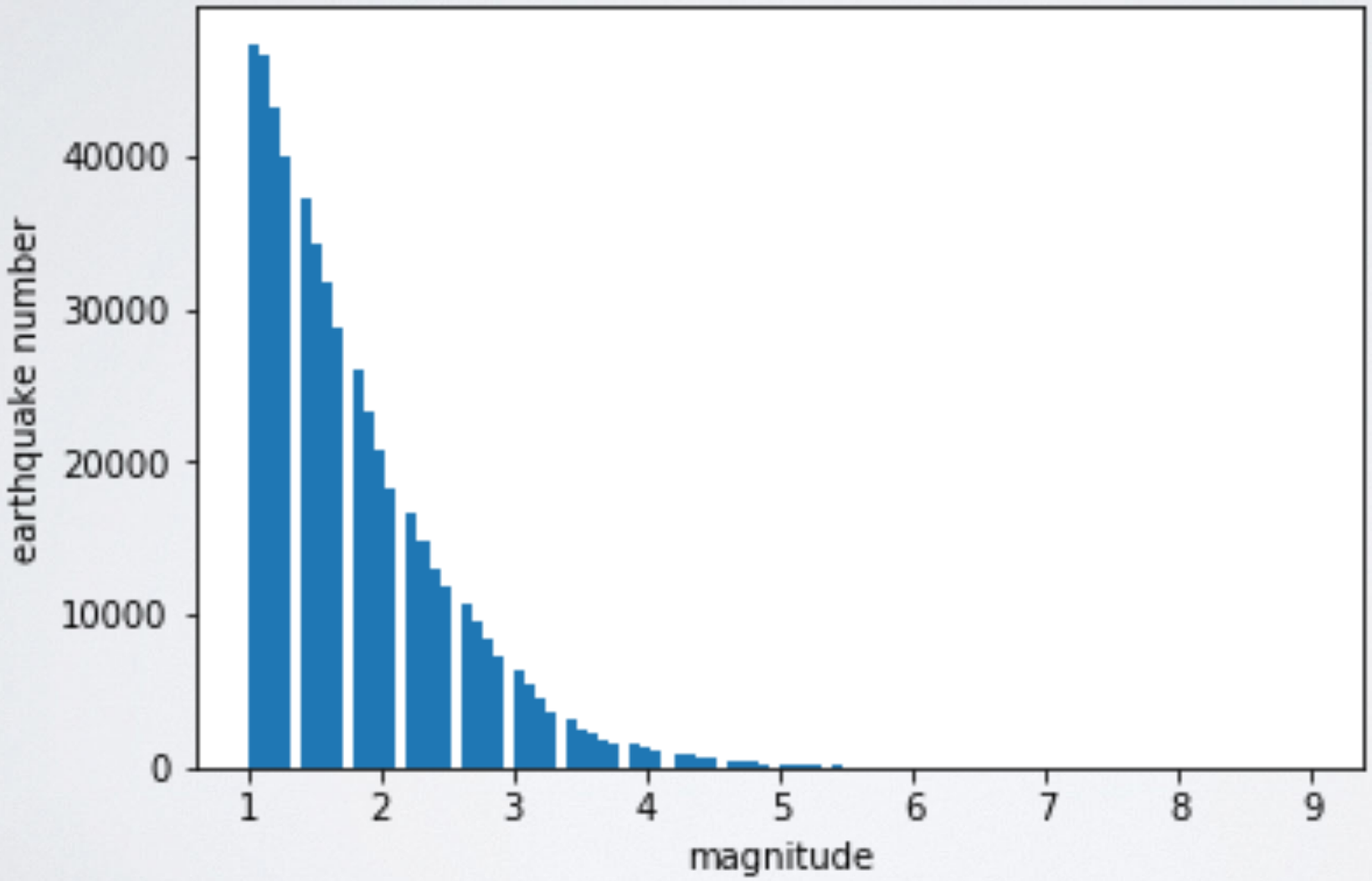


Accuracy:
when:one day
Where: approximate 300km*300km(longitude3 *latitude3)
What: Max (-4)

EARTHQUAKE EVENTS

EARTHQUAKE EVENTS

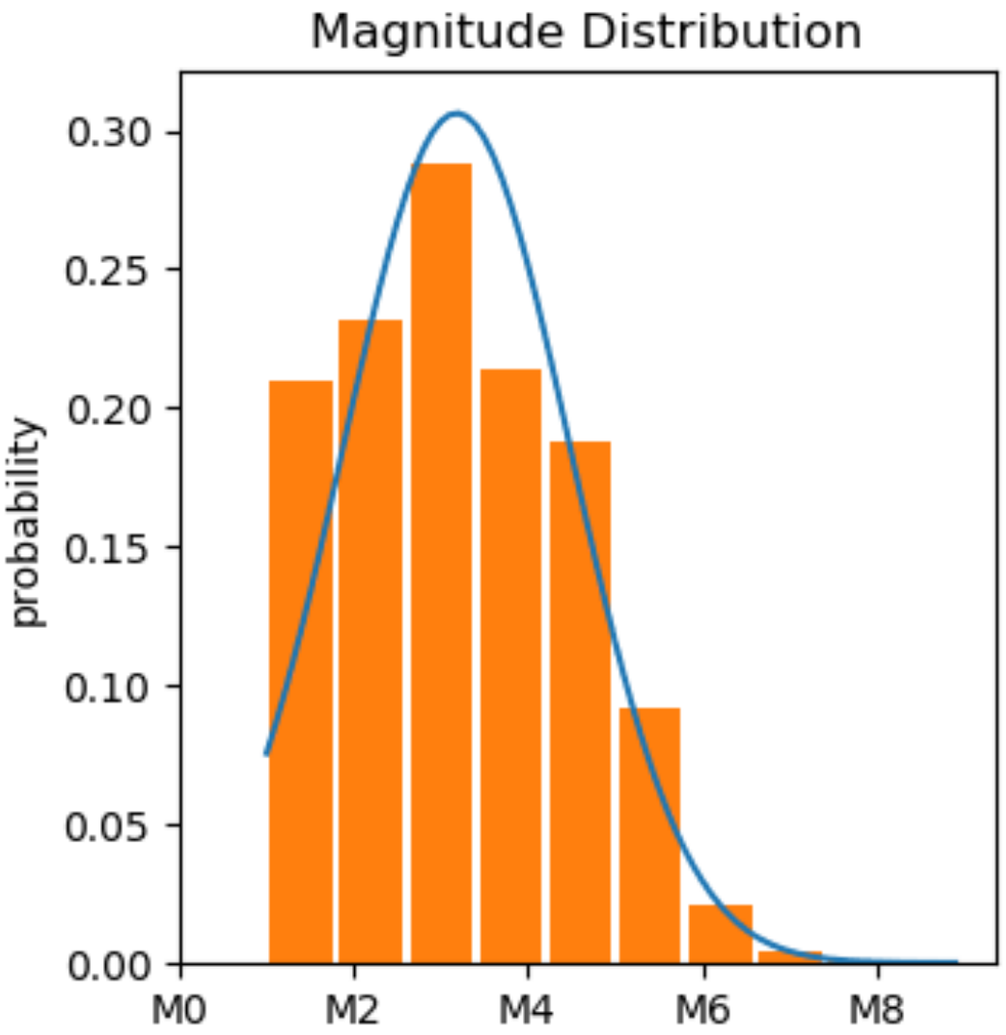
From January 2000-March 2011

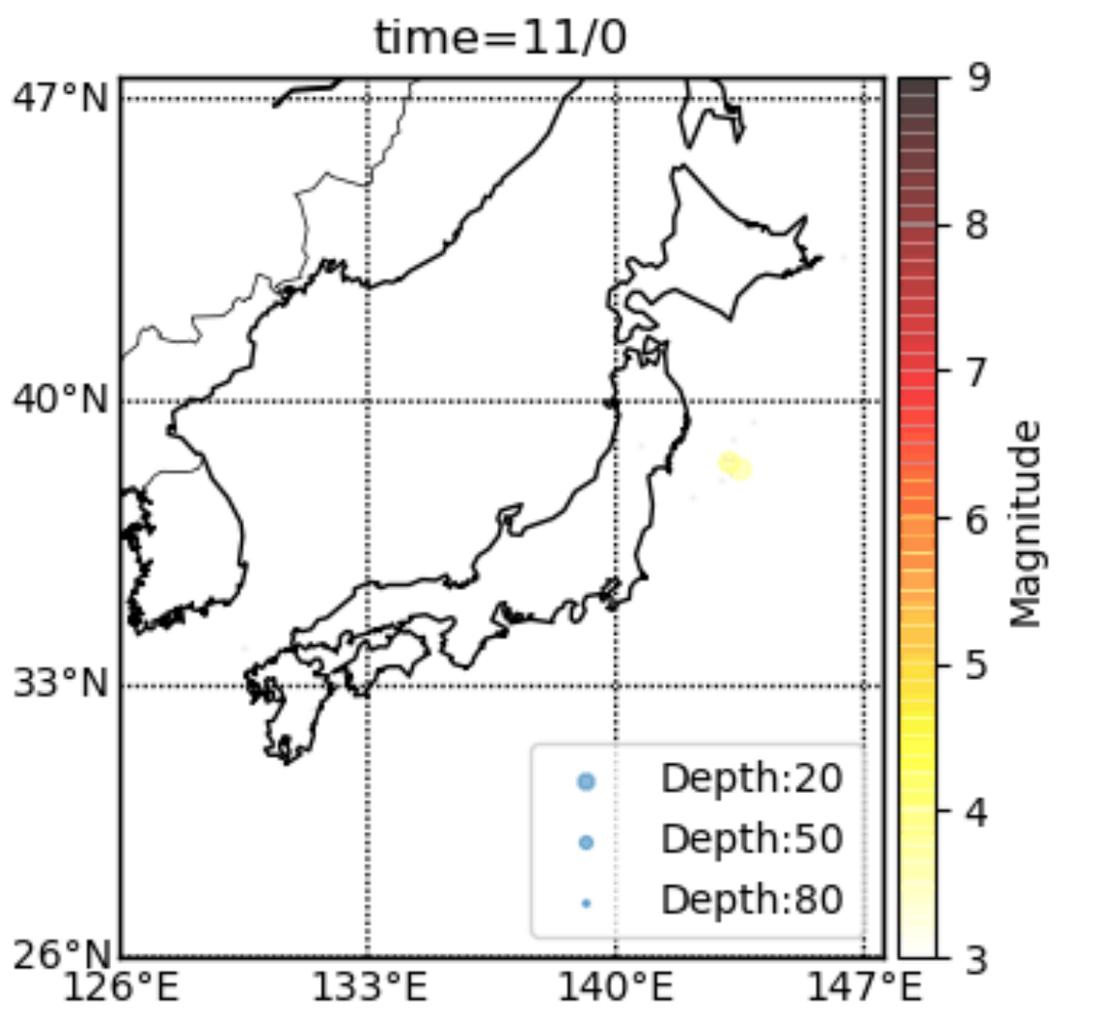


Power law distribution(Continue)

EARTHQUAKE EVENTS

11 March 2011 ⟶ 12 March 2011

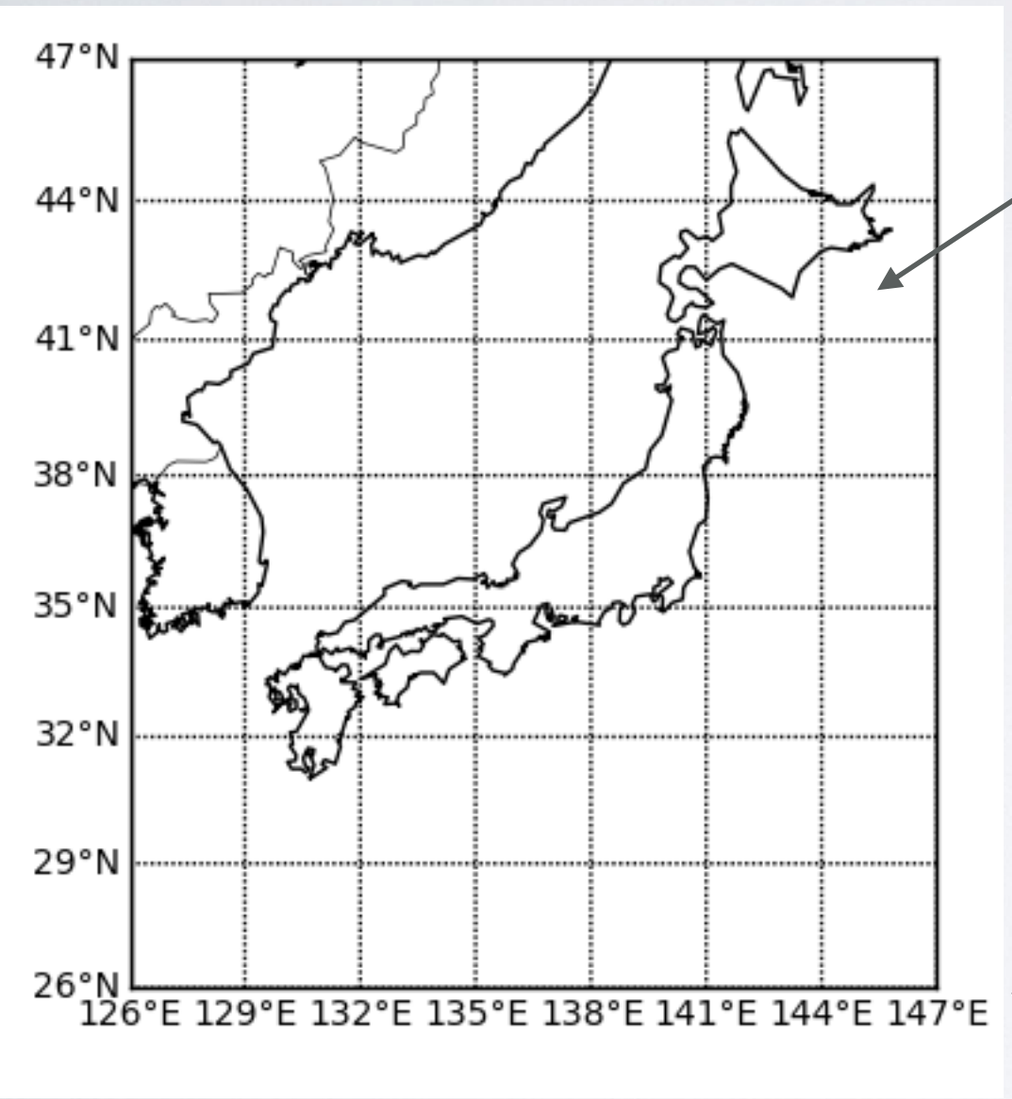




EXPERIMENT

DATA

Miller's cylindrical projection(3D-2D)



Longitude (3 degree),
Latitude(3 degree),
depth<100km

January 2000 ⟶ March 2011

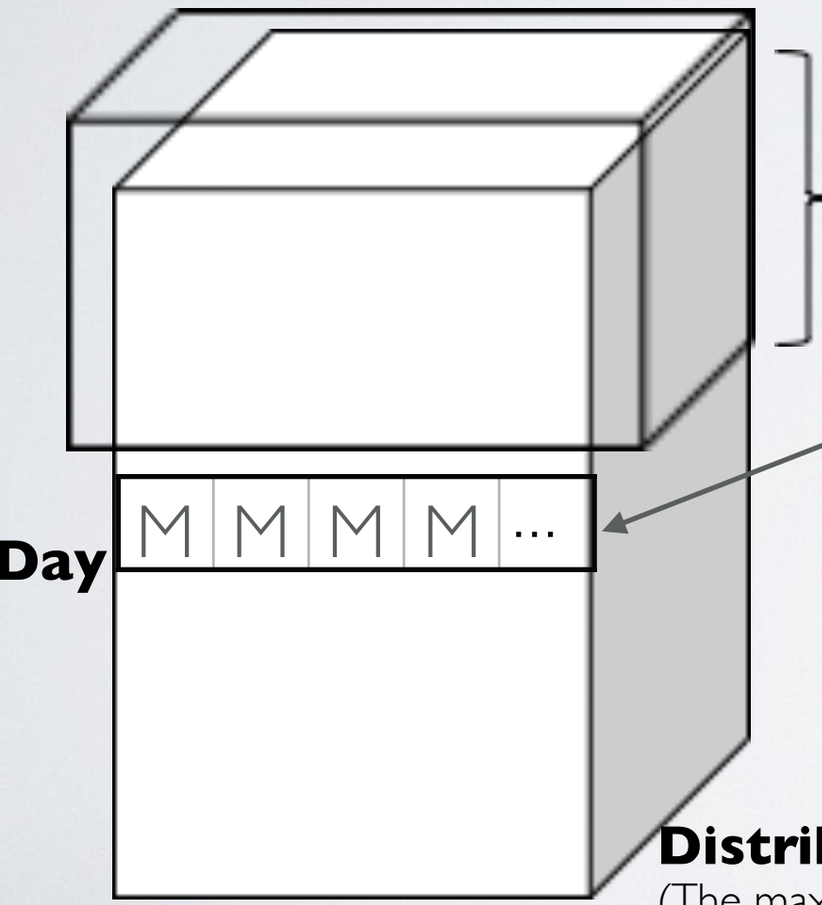
Every day(4087), for every box (49)has a table

One table:
Magnitude distribution value(The maximum value, Median, Mean, Variance, Size(events number))

DATA

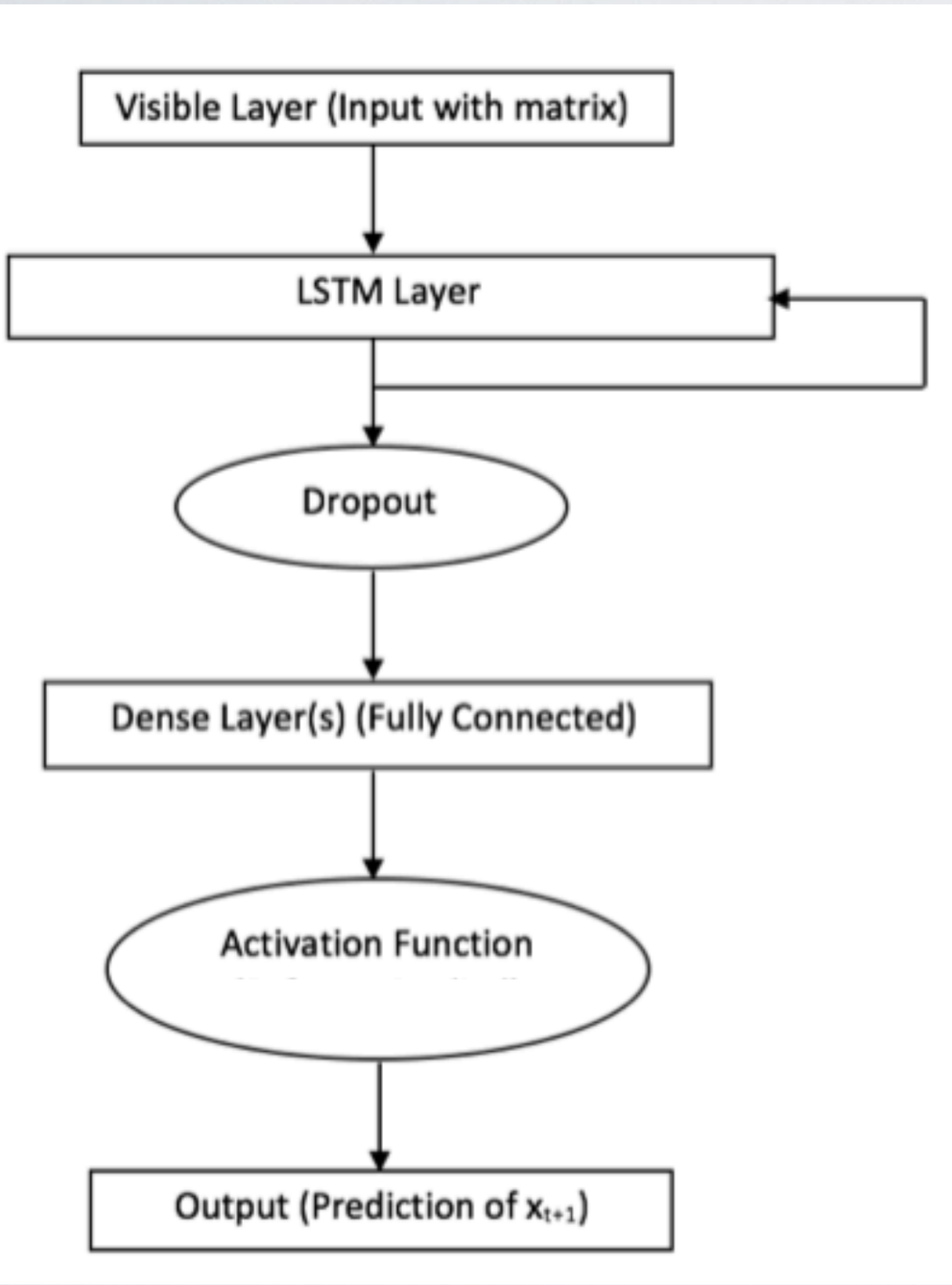
Training datasets: 95% of all datasets(3882*49)
Validation(Testing) datasets: 5% of all datasets(205)
Testing_2 datasets: M>6 in Validation datasets(71)

Training datasets X(Input 3D data):

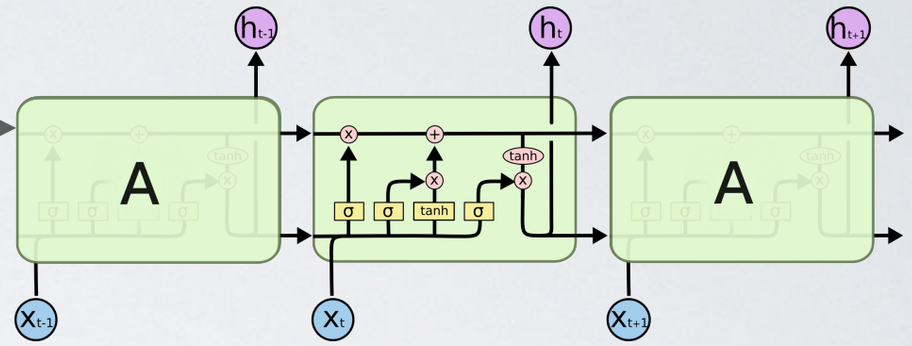


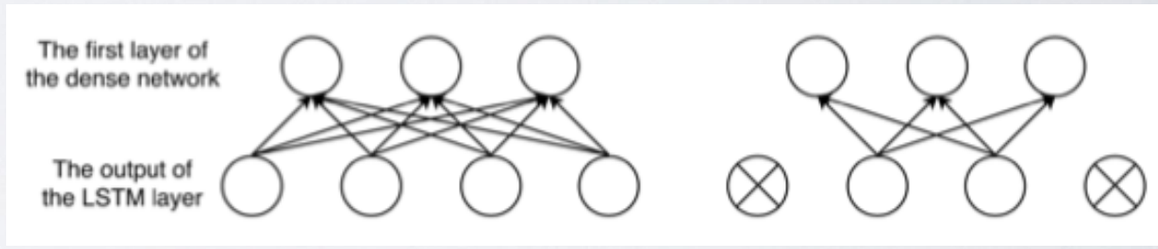
Training datasets Y(Label 2D data):

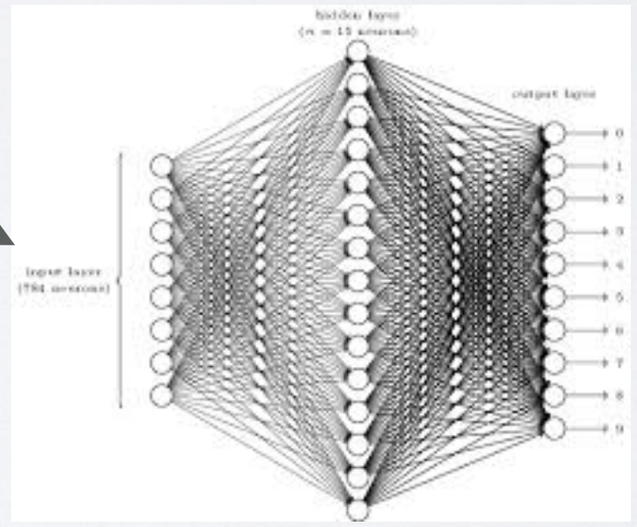
LSTM STRUCTURE



$$\begin{pmatrix} x_{z1}^{t1} \\ x_{z2}^{t1} \\ \vdots \\ x_{z9}^{t1} \end{pmatrix} \begin{pmatrix} x_{z1}^{t2} \\ x_{z2}^{t2} \\ \vdots \\ x_{z9}^{t2} \end{pmatrix} \cdots \begin{pmatrix} x_{z1}^{t11} \\ x_{z2}^{t11} \\ \vdots \\ x_{z9}^{t11} \end{pmatrix}$$







$$S_i = \frac{e^i}{\sum_j e^j}$$

j :Data number

Wang, Qianlong, et al. "Earthquake prediction based on spatio-temporal data mining: an LSTM network approach." IEEE Transactions on Emerging Topics in Computing (2017).

LSTM FOR NLP(NATURAL LANGUAGE PROCESSING)

To be or not to be，that's a question。

Give every man thy ear，but few thy voice。

Take each man's censure，but reserve thy judgement。

—William Shakespeare "Hamlet"

Time window=5 words

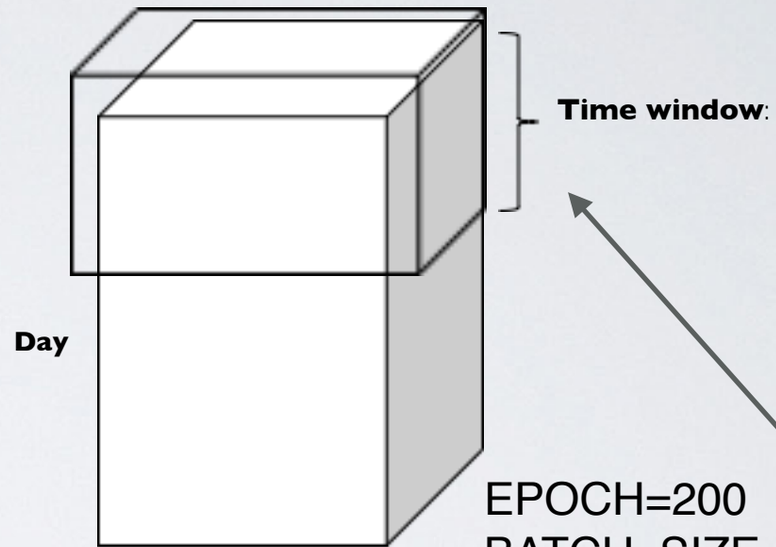
	X	**Y(label)**
Dataset1	"To be or not to"	"be"
Dataset2	"be or not to be"	"that"
⋮	⋮	⋮
Dataset22	"censure but reserve thy"	"judgement"

If the LSTM learn all of the Shakespeare's book, AI will write like Shakespeare.

LSTM STRUCTURE OPTIMIZATION

MSE（Mean squared error）

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$



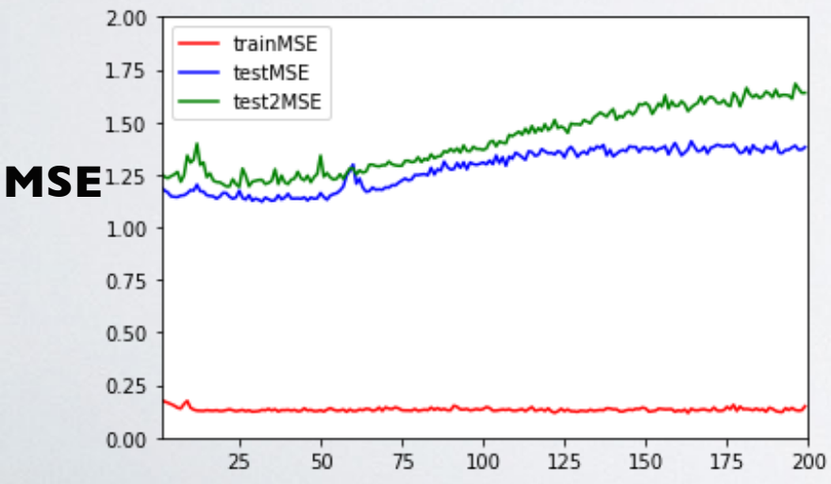
Hyperparameter:

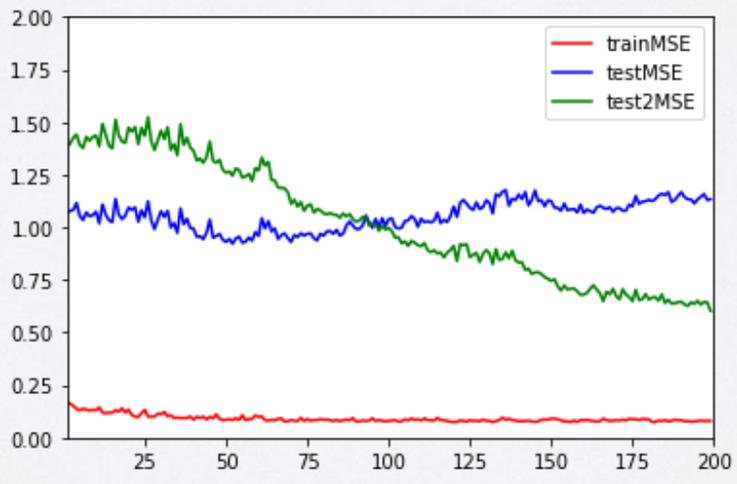
EPOCH=200
BATCH_SIZE=50
TIME_WINDOW=1
LR=0.001

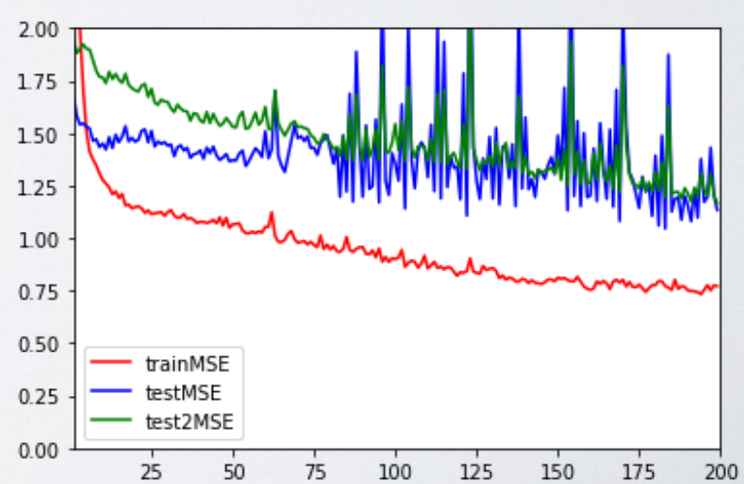
EPOCH=200
BATCH_SIZE=50
TIME_WINDOW=10
LR=0.001

EPOCH=200
BATCH_SIZE=50
TIME_WINDOW=100
LR=0.001

Graph:







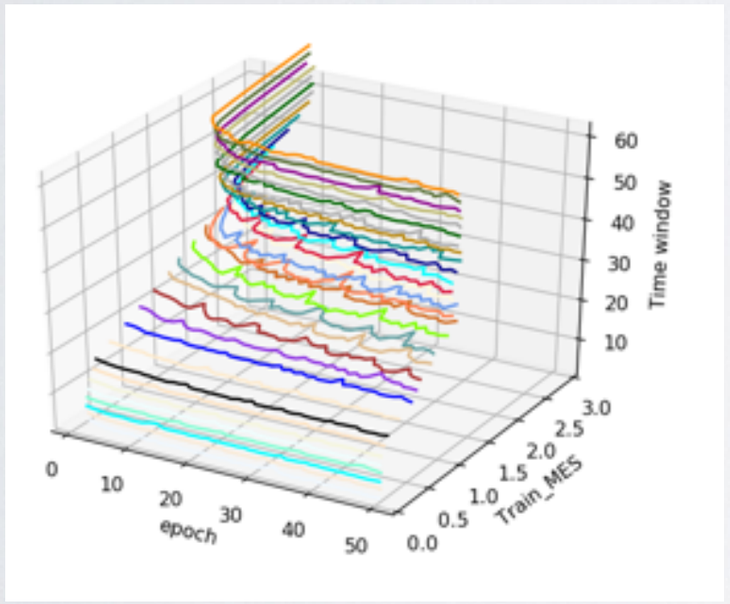
LSTM STRUCTURE OPTIMIZATION

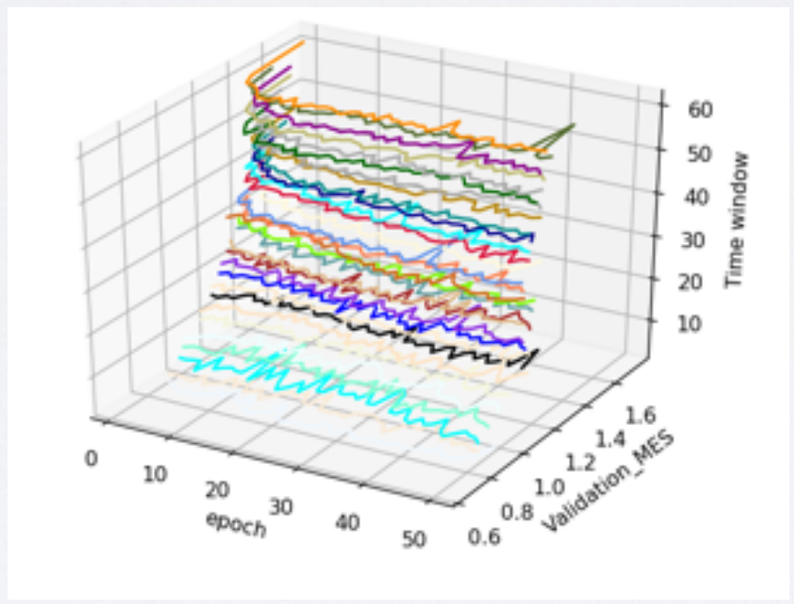
Hyperparameter space:

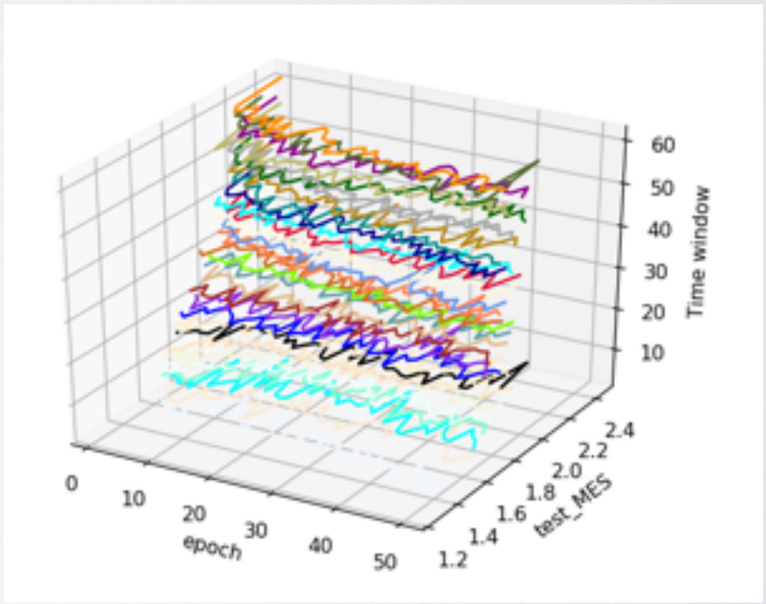
POCH=50
BATCH_SIZE=50
TIME_WINDOW=[0~60]
LR=0.001

POCH=50
BATCH_SIZE=50
TIME_WINDOW=[0~60]
LR=0.001

POCH=50
BATCH_SIZE=50
TIME_WINDOW=[0~60]
LR=0.001







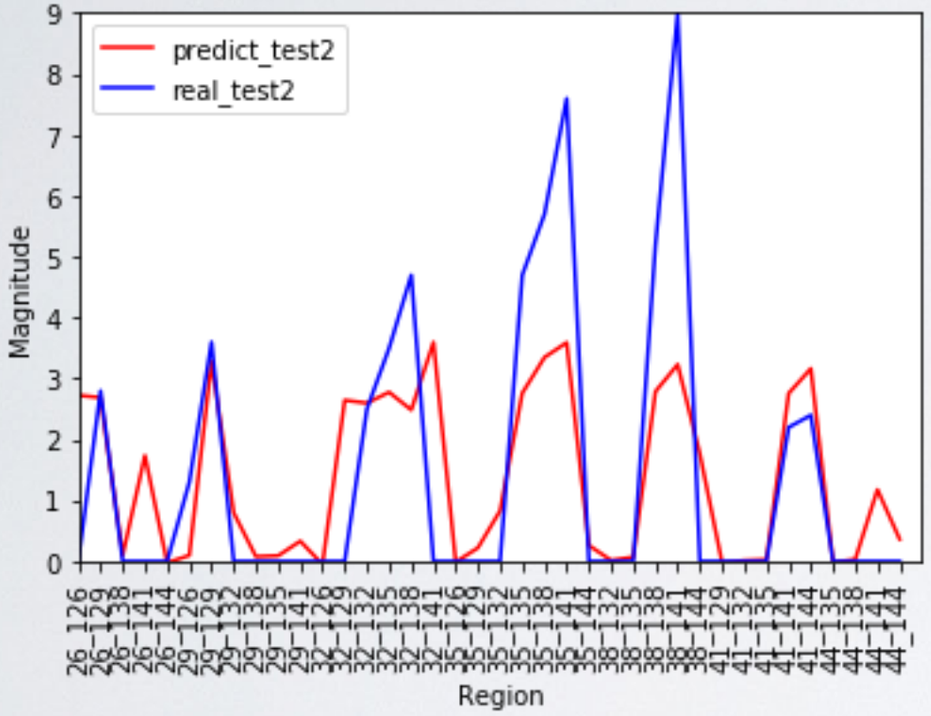
Training

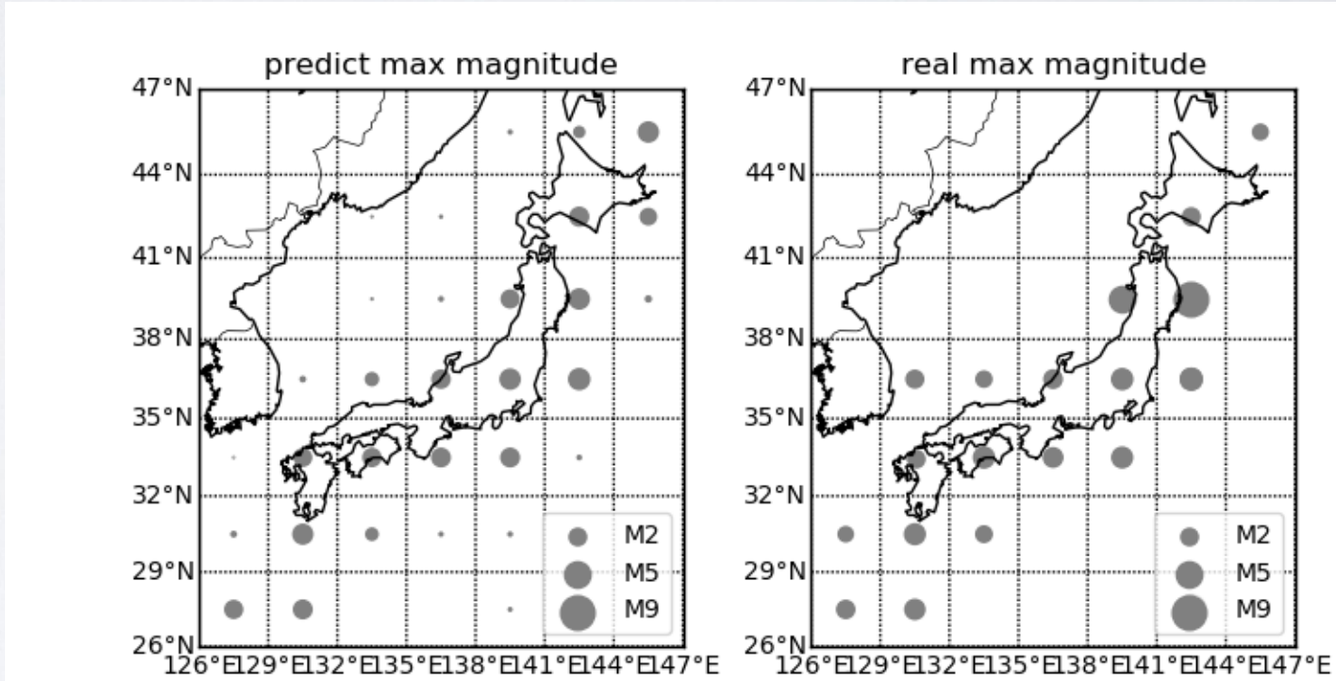
Validation

Testing

LSTM RESULT

2011/3/11

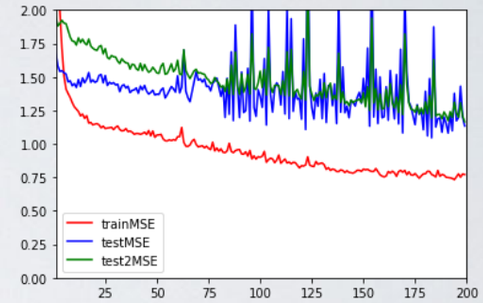




ANALYSIS

Problem:

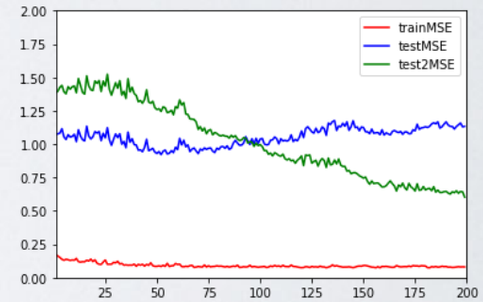
- 1:Validation MSE divergence when training MSE converges(time window <=30)



- 2:Validation variance is large when the Validation MSE converges.(time window >=50)

Result:

- 1:Testing_2 MSE diverge with training MSE converge (time window=100)



- **2: Giant earthquake has a <u>same trend</u> of distribution with all before earthquake as time series data in LSTM.**

EXPECT

- **Optimize the model in Hyper-parameter space**

X=(Batch size, Time window, LR,Epoch)

Y=Validation datasets MSE(30 epoch MSE average after model convergence)

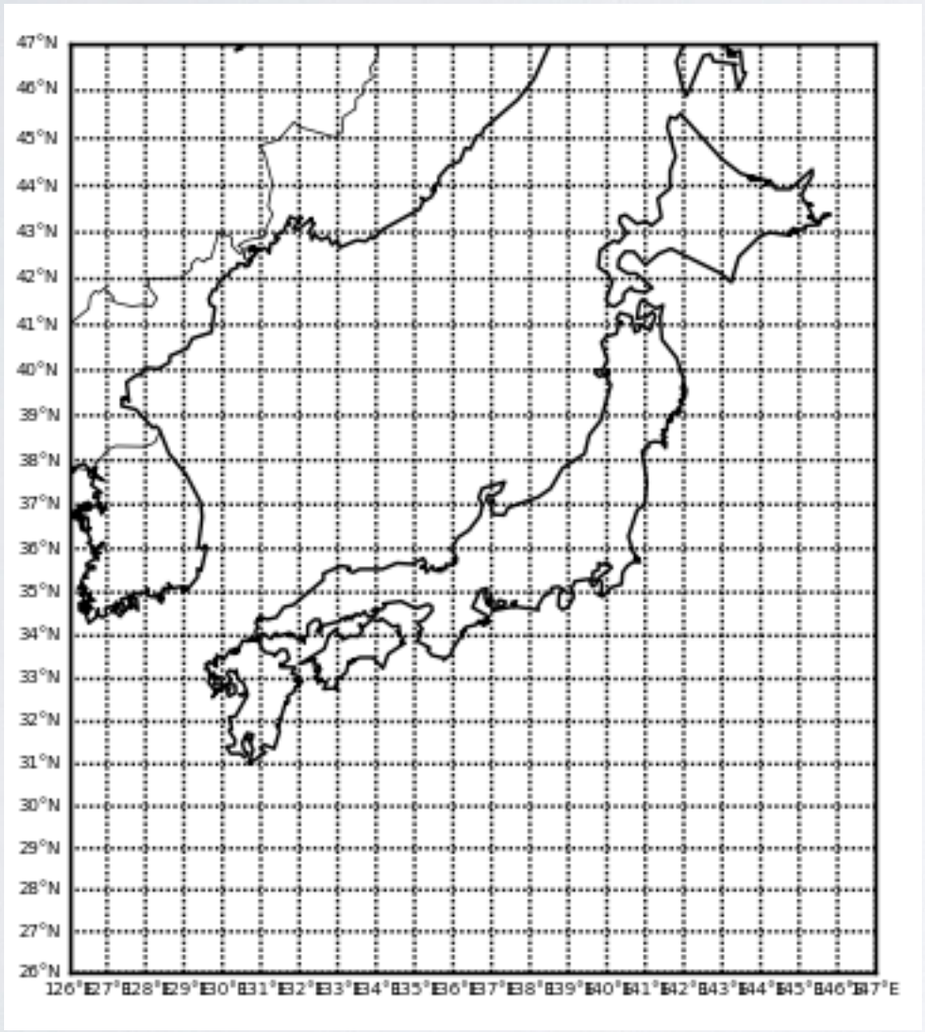
a=(XY)^t(transpose)

New datasets of hyper parameter:A:{a|(XY)^t}

In A space find optimized point **a**.
And find hyper-parameter's law from MDS(Multidimensional scaling)

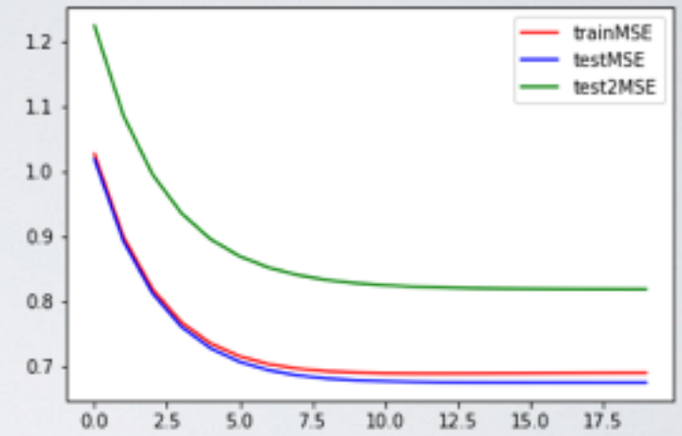
EXPECT

After we got optimize model:

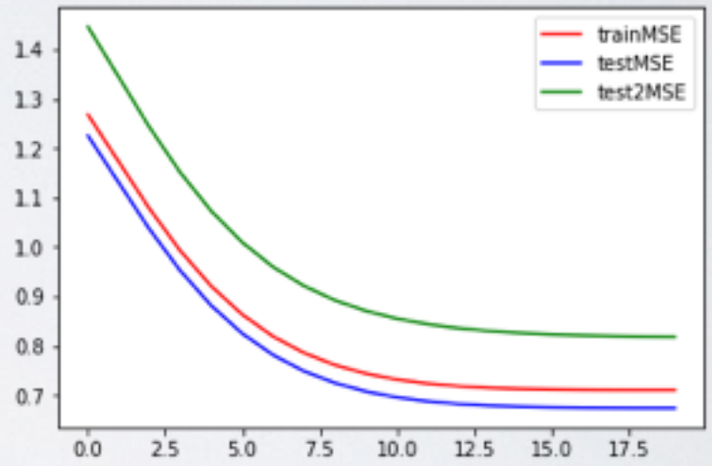


Space scale -> 1degree

Epoch=20
BATCH_SIZE=50
TIME_WINDOW=10
LR=0.005



Epoch=20
BATCH_SIZE=50
TIME_WINDOW=30
LR=0.005



Optimize

Epoch=100
BATCH_SIZE=50
TIME_WINDOW=?
LR=?

?

SUMMARY

REFERENCES

- http://colah.github.io/posts/2015-08Understanding-LSTMs/
- http://karpathy.github.io/2015/05/21/rnn-effectiveness/
- Li Z, Meier M A, Hauksson E, et al. Machine Learning Seismic Wave Discrimination: Application to Earthquake Early Warning[J]. Geophysical Research Letters, 2018.
- Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.
- https://wwweic.eri.u-tokyo.ac.jp/db/jma.deck/index-j.html
- Wang, Qianlong, et al. "Earthquake prediction based on spatio-temporal data mining: an LSTM network approach." IEEE Transactions on Emerging Topics in Computing (2017).