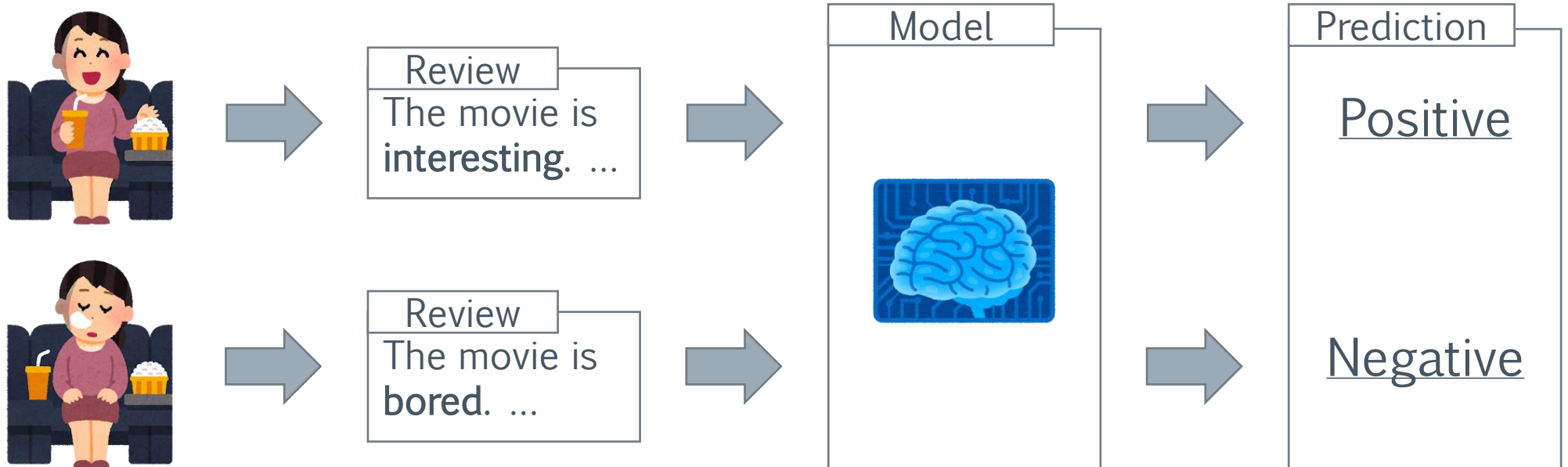


Comparison of classification accuracy between different pattern recognition methods

Okai Seiyu, IST
2IE21093N Okai Seiyu

Purpose of Exercise

- › IMDB Datasets(Movie Review Datasets)
- › Classify **positive** review or **negative** review.
- › Train data 25000,Test data 25000



- › One of my seniors in the lab was conducting a similar study, and my research is related to NLP.

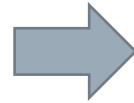
Approach

› I used four pattern recognition methods

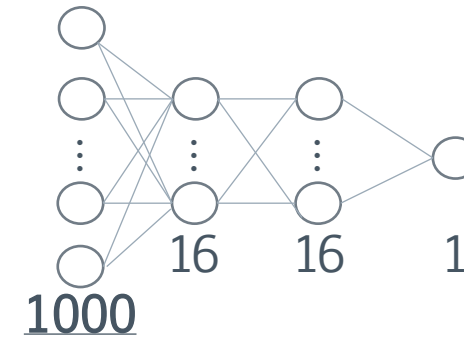
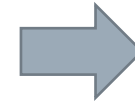
– Deep Neural Network

Use multi-hot vector

Review
The movie is
interesting. ...



$[0, 1, 1, \dots]^T$
1000 dim



– LSTM

› A special kind of RNN

– Naive Bayes Classifier

› Use Bernoulli distribution

– Light GBM

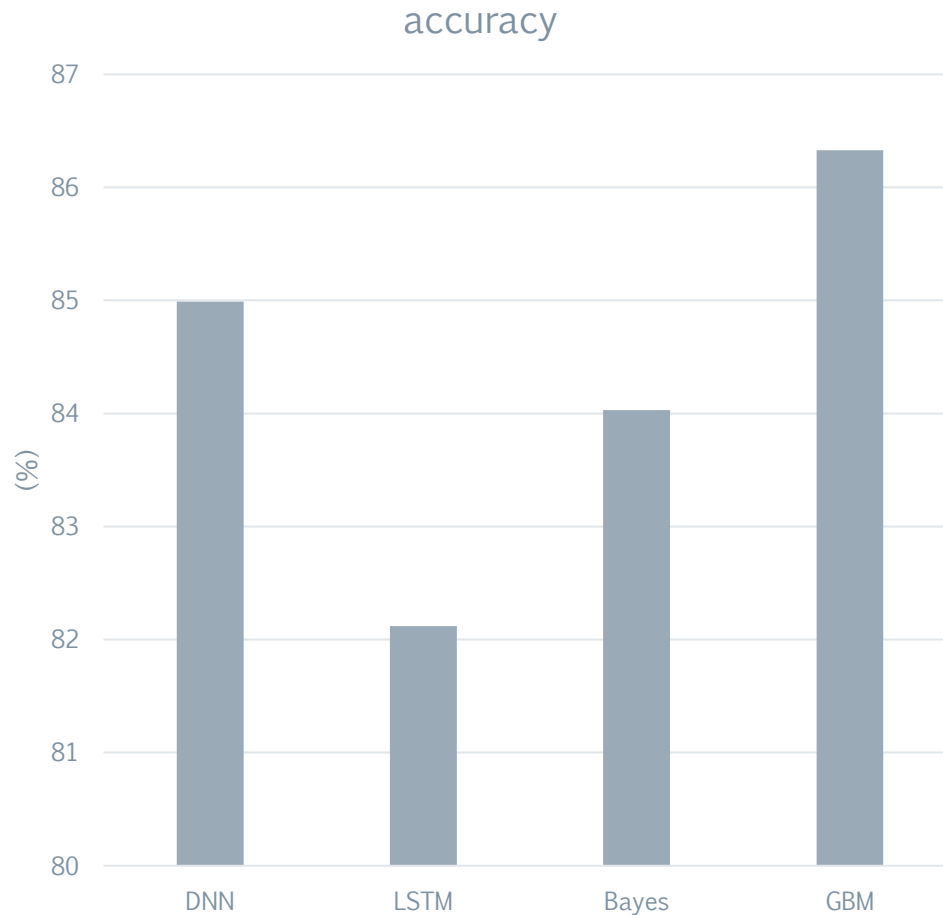
› Ensemble Learning

Use context information

Use distribution model

Use decision tree

Results and Conclusion



- › Light GBM was the best method of the four methods.
- › LSTM had the lowest accuracy, which was surprising.
- › However, calculation time and study time are not taken into account in the evaluation.
- › A more comprehensive evaluation should be conducted.