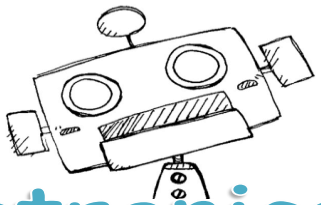
National Taiwan University
Department of Bio-industrial Mechatronics Engineering
Bio-mechatronics Lab

Titanic

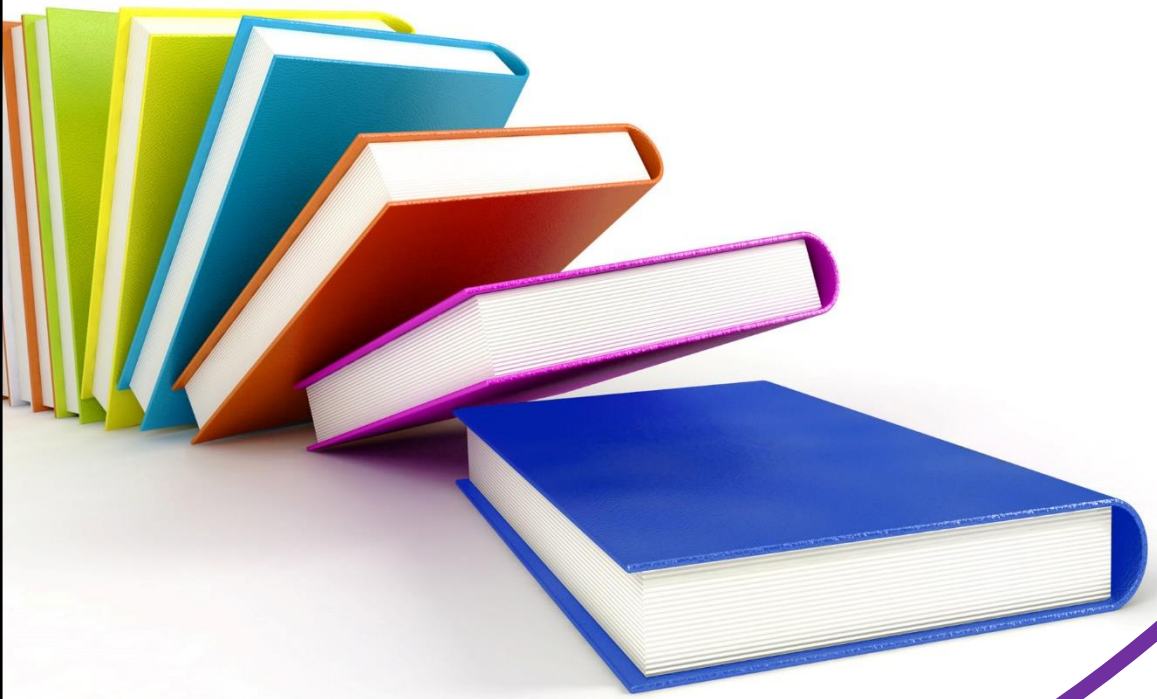
Author: Warren Tseng



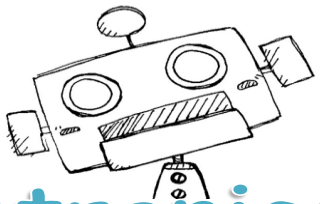
Biomechatronics

Outline

Data Pre-processing	Training Model	Result
- Feature Selection - Feature Processing	- NN Structure - Hyper-parameter	



Data Processing



Biomechatronics

Feature Selection

- **Pclass, fare**: social stratification
- **Sex**: lady first
- **Age**: baby first
- **Sibsp, parch**: they are numeric value…

Feature Processing

- **Pclass, sex: from class to ‘one hot’ value.**

1	**→**	**[1, 0, 0]**	**male**	**→**	**[1, 0]**
2	**→**	**[0, 1, 0]**	**female**	**→**	**[0, 1]**
3	**→**	**[0, 0, 1]**			

Feature Processing

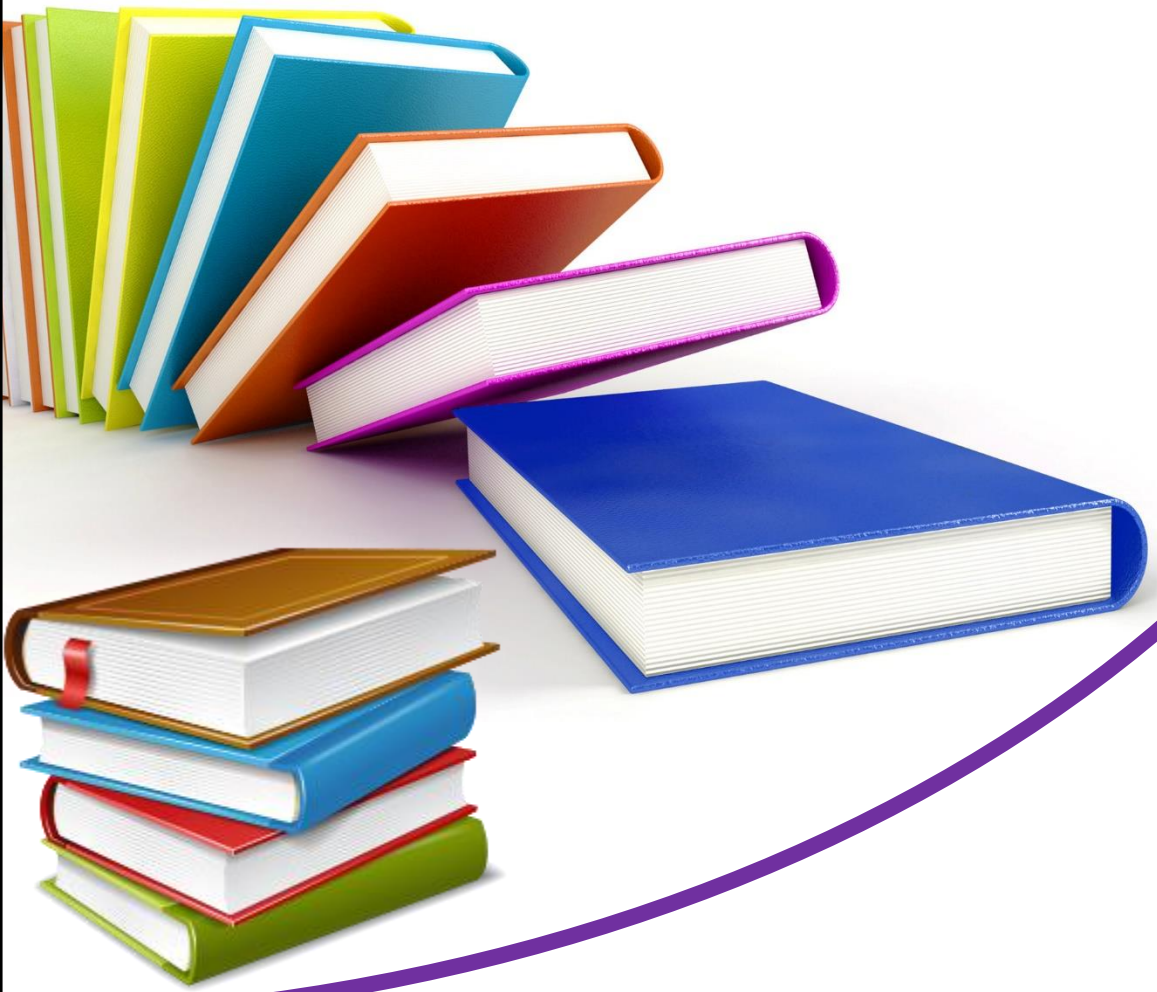
- **Age: missing value → -100, normalization → age / 100**
- **Fare: missing value → -200, normalization → fare / 200**
- **Sibsp: normalization → sibsp / 200**
- **Parch: normalization → Parch / 200**

Feature Processing

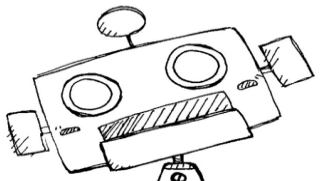
- **Survived: from class to ‘one hot’ value.**

0 → [1, 0]

1 → [0, 1]

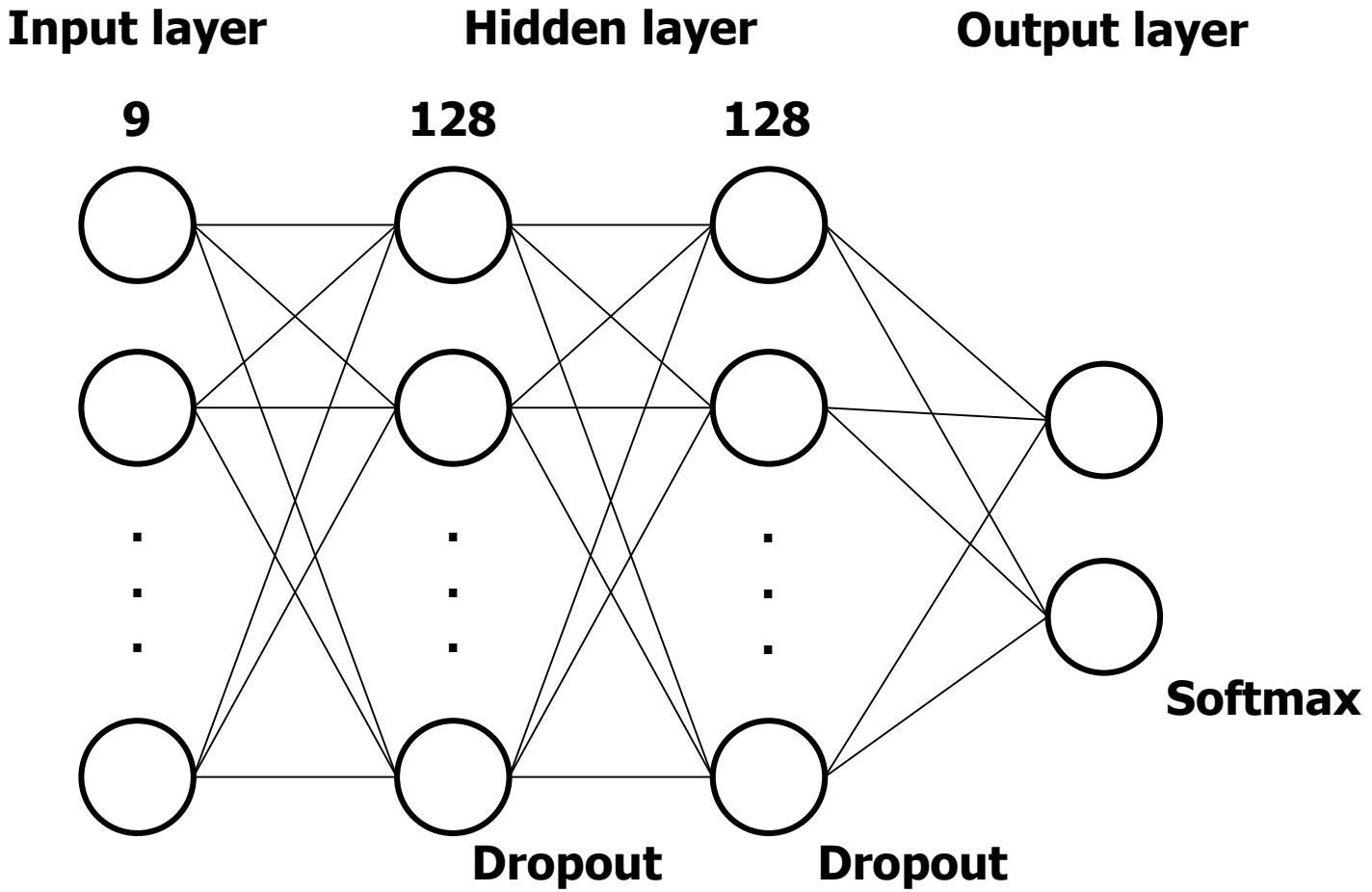


Training Model



Biomechatronics

NN Structure



Hyper-parameters

- Epoch = 2000
- Learning rate = 1e-4
- Dropout probability = 0.2
- Loss function: cross-entropy
- Optimizer: RMSprop

Result

```
Epoch: 100 , LossTrain: [ 0.542] , AccuTrain: 0.79125 , AccuTest: 0.78391959799
Epoch: 200 , LossTrain: [ 0.517] , AccuTrain: 0.7925 , AccuTest: 0.798994974874
Epoch: 300 , LossTrain: [ 0.509] , AccuTrain: 0.80375 , AccuTest: 0.793969849246
Epoch: 400 , LossTrain: [ 0.505] , AccuTrain: 0.79875 , AccuTest: 0.798994974874
Epoch: 500 , LossTrain: [ 0.502] , AccuTrain: 0.8 , AccuTest: 0.819095477387
Epoch: 600 , LossTrain: [ 0.5] , AccuTrain: 0.80375 , AccuTest: 0.814070351759
Epoch: 700 , LossTrain: [ 0.498] , AccuTrain: 0.815 , AccuTest: 0.819095477387
Epoch: 800 , LossTrain: [ 0.493] , AccuTrain: 0.81375 , AccuTest: 0.814070351759
Epoch: 900 , LossTrain: [ 0.488] , AccuTrain: 0.8275 , AccuTest: 0.809045226131
Epoch: 1000 , LossTrain: [ 0.485] , AccuTrain: 0.82875 , AccuTest: 0.804020100503
Epoch: 1100 , LossTrain: [ 0.483] , AccuTrain: 0.82625 , AccuTest: 0.804020100503
Epoch: 1200 , LossTrain: [ 0.481] , AccuTrain: 0.82625 , AccuTest: 0.804020100503
Epoch: 1300 , LossTrain: [ 0.48] , AccuTrain: 0.825 , AccuTest: 0.793969849246
Epoch: 1400 , LossTrain: [ 0.478] , AccuTrain: 0.8275 , AccuTest: 0.793969849246
Epoch: 1500 , LossTrain: [ 0.477] , AccuTrain: 0.8275 , AccuTest: 0.788944723618
Epoch: 1600 , LossTrain: [ 0.477] , AccuTrain: 0.82875 , AccuTest: 0.793969849246
Epoch: 1700 , LossTrain: [ 0.476] , AccuTrain: 0.83125 , AccuTest: 0.793969849246
Epoch: 1800 , LossTrain: [ 0.475] , AccuTrain: 0.83 , AccuTest: 0.788944723618
Epoch: 1900 , LossTrain: [ 0.475] , AccuTrain: 0.83125 , AccuTest: 0.788944723618
Epoch: 2000 , LossTrain: [ 0.474] , AccuTrain: 0.83125 , AccuTest: 0.788944723618
```

Conclusion

- **Overfitting for 2000 epochs**

- **The overfitting was occurred when the epoch was from 800 to 900.**

  → 1. Epoch ↓

  2. Dropout probability ↑

Thanks for your attention