Group 10

Chung Yin Chan
Sangin Ha
Pok Lam Tang
Shuk Hung Li

Project 1 sentiment analysis - baseline model

LSTM

- Batch normalization
- N-gram

Logistic Regression

- C = 10

Project 1 sentiment analysis - BERT Model

- Batch size = 16
- Learning rate = 1e-5
- Result:

```
Epoch 1
Training loss: 0.8637171213506829
Validation loss: 0.7545768254626114
F1 Score (Weighted): 0.6795930920000809

Epoch 2
Training loss: 0.654190214312206
Validation loss: 0.7300146029586285
F1 Score (Weighted): 0.699584630099294

Epoch 3
Training loss: 0.5616043685895905
Validation loss: 0.7652329509764646
F1 Score (Weighted): 0.6933984084558836
```
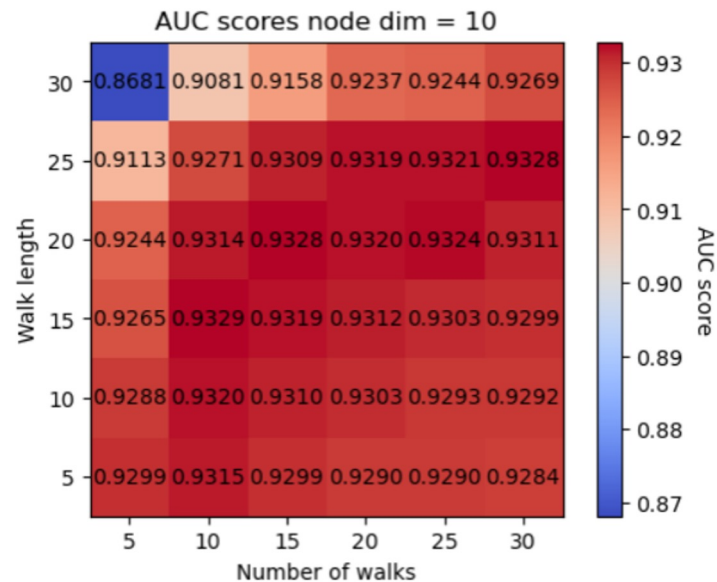
project 2 social network mining - Model comparison

	DeepWalk	**Node2Vec**
Hyperparameter tested	# of walks - Walk length - Node dimension (5-30, increment of 5)	# of walks - Walk length - P and Q values
Best AUC score (# of walks = 10, walk length = 20)	0.9329	0.9317

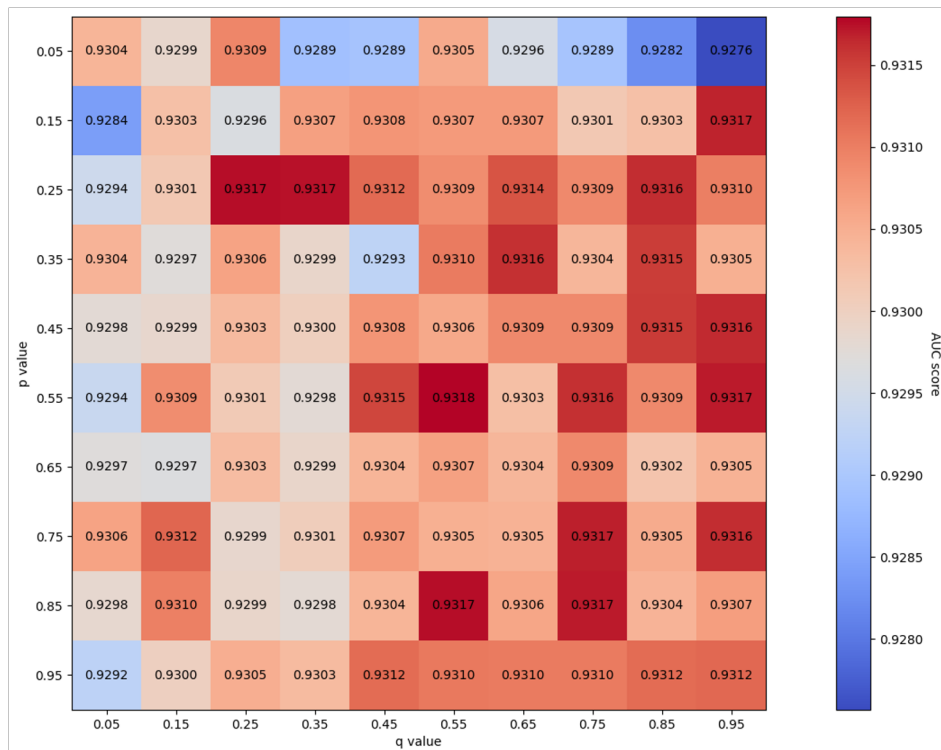# project 2 social network mining - Deepwalk model

Optimal hyperparameter values
- node_dim=10, num_walks=10, walk_length=15
- p=0.5, q=0.5
- AUC score of 0.9329

AUC scores node dim = 10

PROJECT 2 SOCIAL NETWORK MINING - NODE2VEC MODEL

Optimal hyperparameter values
- node_dim=10, num_walks=10, walk_length=20
- p=0.55, q=0.55
- AUC score of 0.9318

Q&A