



# STAT 5243 - Project 3: Weakly Supervised Learning

## Group 9:

Yu, Yang  
Wang, Jingqi  
Lee, Rhoan  
Zhou, Jitian

# Introduction

- Baseline model
- Advanced models (Model I & Model II) by CNN



# Baseline Model

- RGB histogram features
- Noisy labels
- Logistic regression model



# Model I

- More sophisticated model than the baseline model
- CNN model  
(using **TensorFlow** and **Keras**)
- Treat noisy labels as clean labels



# Model II

- The same predictive model as in Model I
- CNN model  
(using **TensorFlow** and **Keras**)
- But address the label noise issue



# Proposal

- 01 Performance improvement
- 02 Running cost tradeoff
- 03 Supporting evidence



# Proposal - Summary

Performance Improvement & Running Cost Tradeoff

	Baseline Model	Model I	Model II
Accuracy	0.24	0.49	0.57
Time	12 s	7 min	13 min
Memory (MB)	0.00	124.16	259.60



# Proposal - Strategy & Supporting Evidence

- Strategy:

- Model II is recommended
- Mitigate the running costs

- Supporting Evidence:

- The tradeoff in running costs is justified by the considerable gains in prediction accuracy





Thank you!

