1. Loss function vs floating point operations (varying batch size & constant learning rate):

Batch Size: 800, Learning Rate: 1

Batch Size: 1000, Learning Rate: 1

Cross Entropy Loss

Floating points operations/ n*m*k
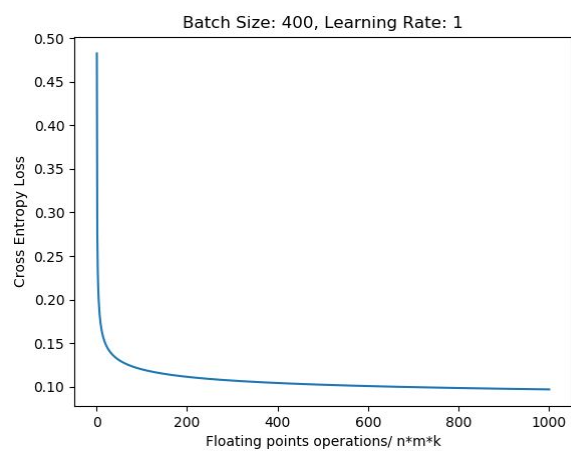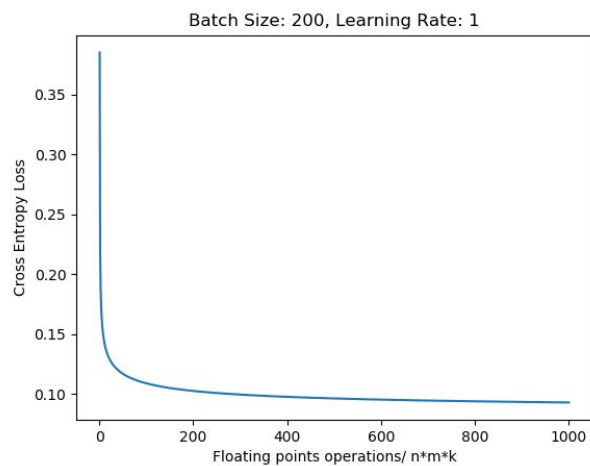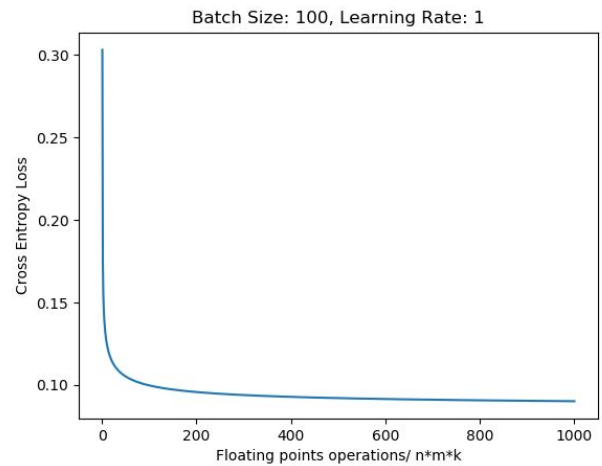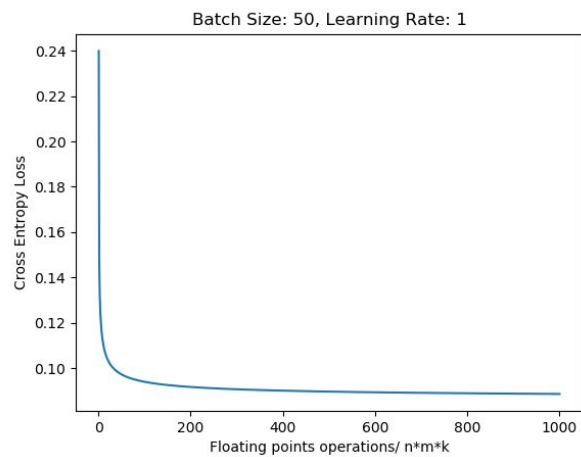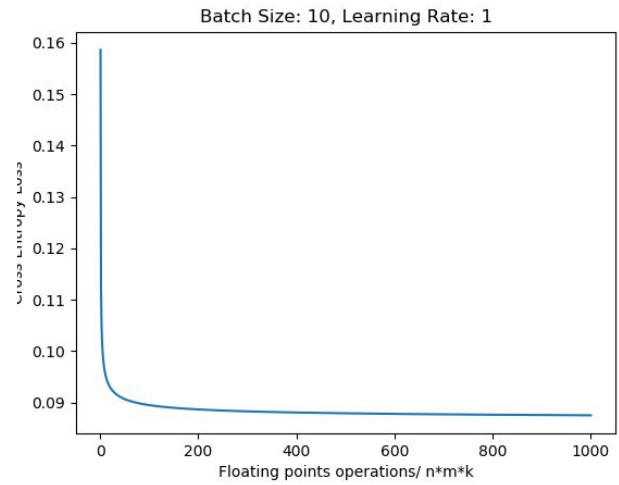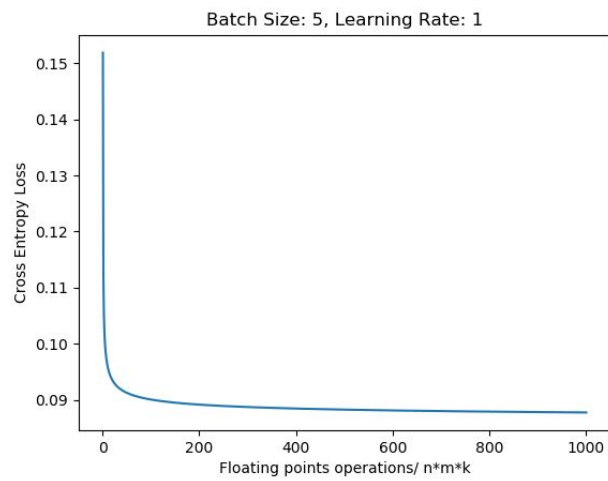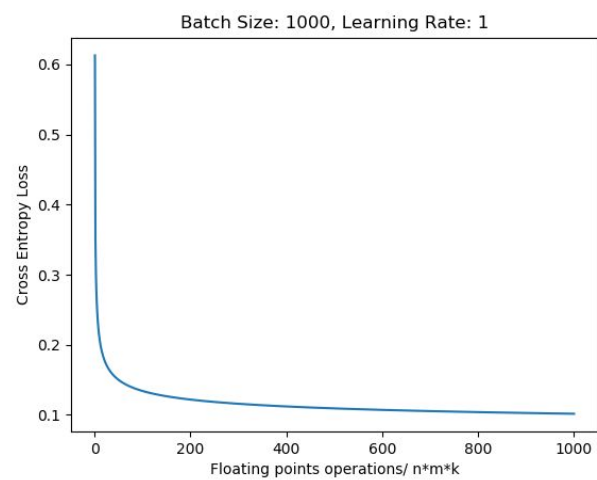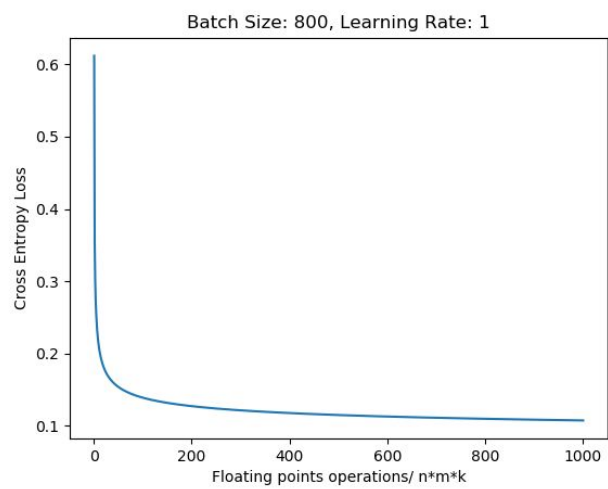
2. Loss function vs floating point operations (varying learning rate & constant batch size):
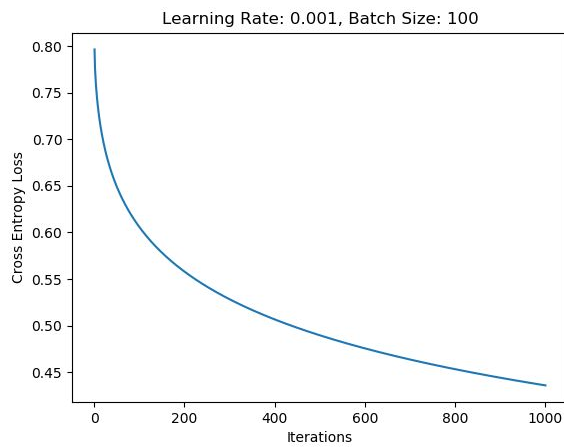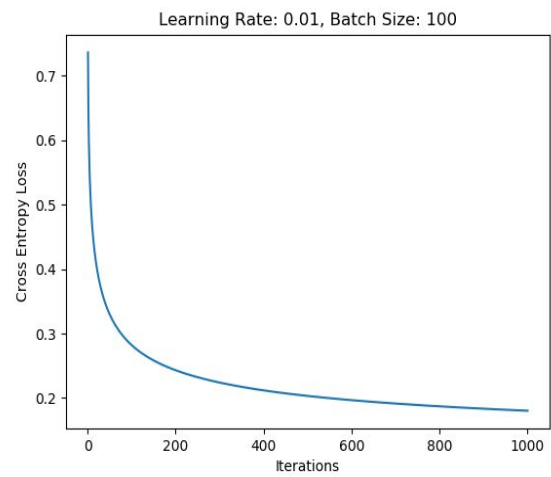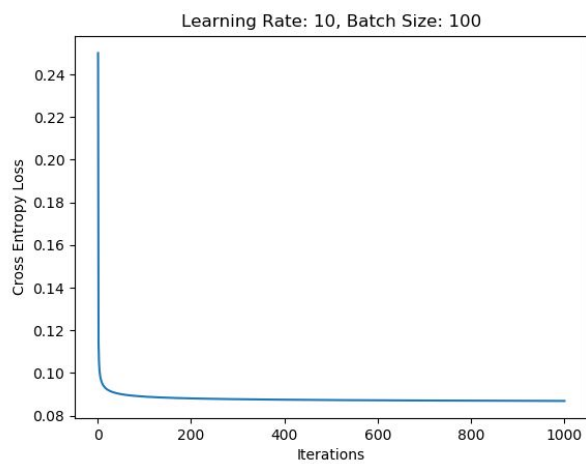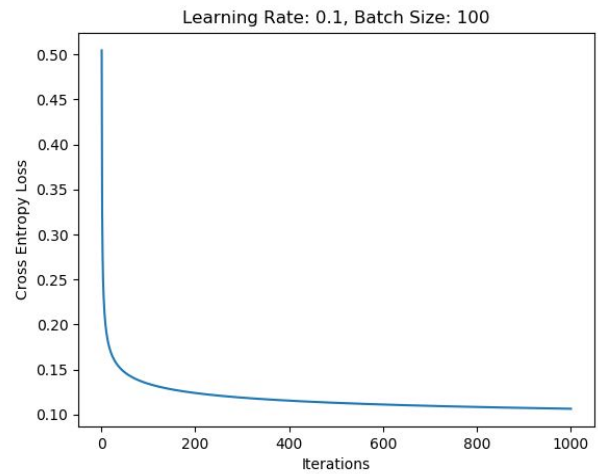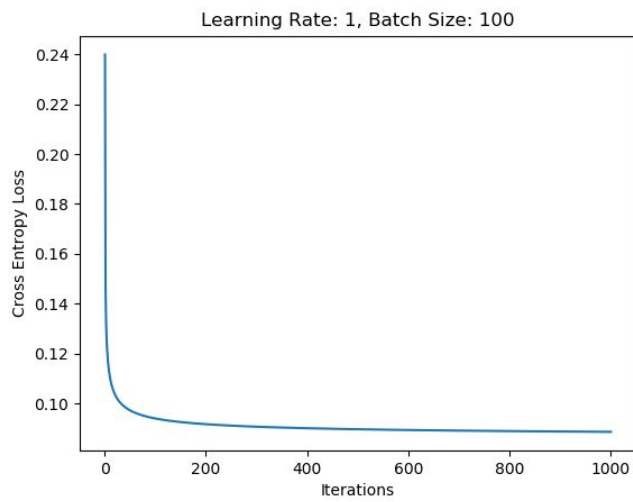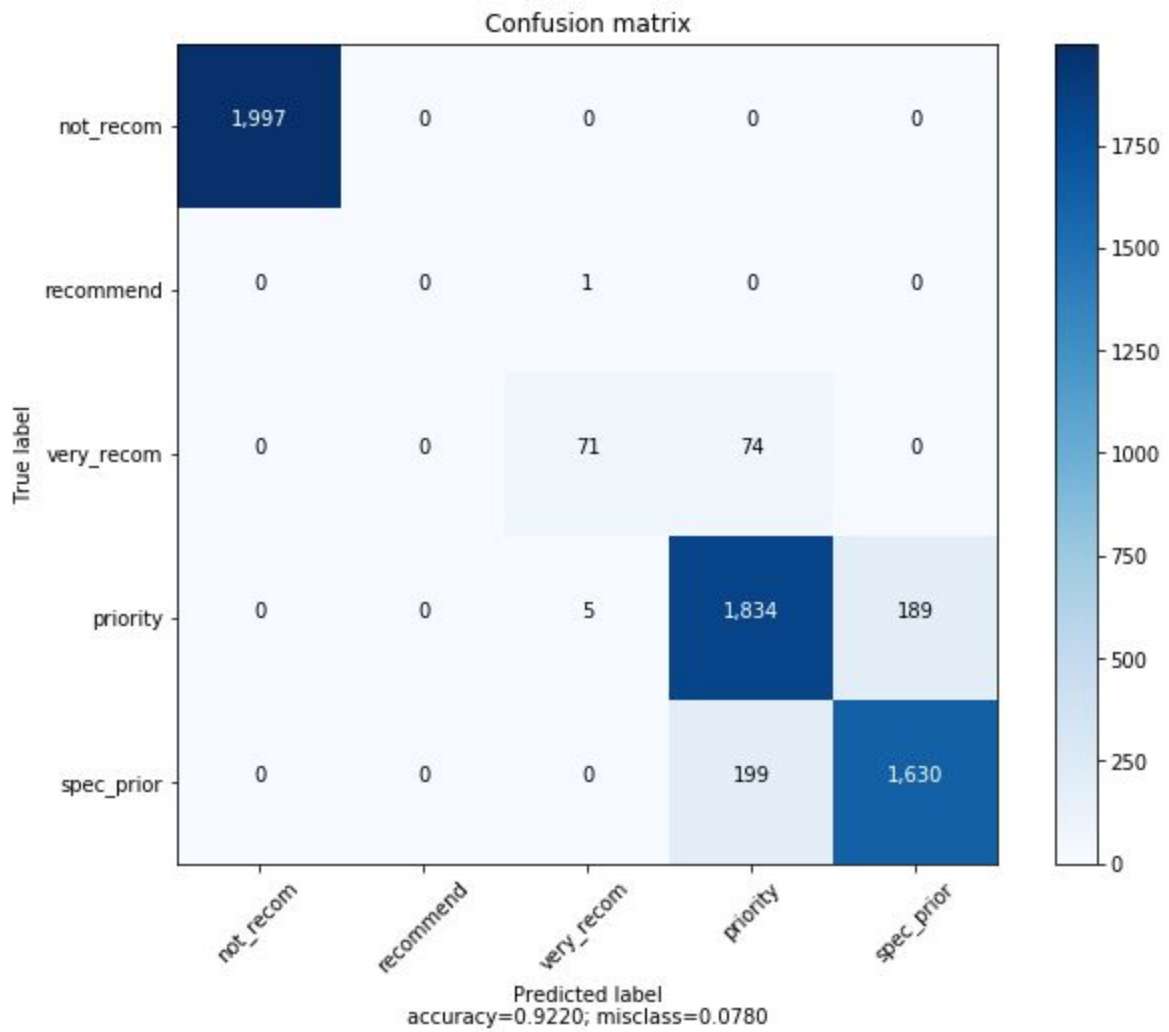

** X-axis represents no. of floating point operations/ n*m*k

3. Confusion Matrix:

F1 Scores:

Class Name:	not_recom	recommended	very_recom	priority	spec_prior
F1_Score	1.00	0.00	0.64	0.89	0.89

Macro F1 Score: 0.684
Micro F1 Score: 0.922