Dataset Split:

70% training
10% validation
20% testing

Feature Selection:

Top N most frequent words (stop words excluded) in the dataset
N = 2500

Results:

KNN: K = 3

Classifier	Without Feature Selection				
	Training	Validation	Testing	Offline Efficiency Cost	Online Efficiency Cost
Decision Tree	99.71%	69.67%	72.0%	1853 ms	7 us
KNN	84.9%	74.67%	75.5%	33 us	20 ms

Classifier	With Feature Selection				
	Training	Validation	Testing	Offline Efficiency Cost	Online Efficiency Cost
Decision Tree	99.38%	73.67%	74.33%	799 ms	3 us
KNN	86.95%	75.33%	78.0%	22 us	14 ms

Feature selection improves both online and offline efficiency cost

Decision Tree		
	Offline Efficiency Cost	Online Efficiency Cost
Without Feature Selection	1853 ms	7 us
With Feature Selection	799 ms	3 us

KNN		
	Offline Efficiency Cost	Online Efficiency Cost
Without Feature Selection	33 us	20 ms
With Feature Selection	22 us	14 ms

Feature selection improves accuracy on testing dataset

Accuracy on Testing Dataset		
Classifier	Without Feature Selection	With Feature Selection
Decision Tree	72.0%	74.33%
KNN	75.5%	78.0%

KNN has better offline efficiency & Decision tree has better online efficiency

With Feature Selection		
Classifier	Offline Efficiency	Online Efficiency
Decision Tree	1853 ms	7 us
KNN	33 us	20 ms

Without Feature Selection		
Classifier	Offline Efficiency	Online Efficiency
Decision Tree	799 ms	3 us
KNN	22 us	14 ms