

### 3. Training

Grid search for optimal  $C$  configuration

Five-fold  
Cross  
Validation  
- Linear  
Model

Best  $C$  configuration

Final SVM

### 4. Classification

Model type:  
L2 regularized logistic regression

Final  
Linear  
SVM

### 5. Results

Raw Prediction Output

PubMed Id	Prediction Set	Rank A	Rank B
PMID1	A	0.98	0.02
PMID2	A	0.51	0.49
PMID3	B	0.02	0.98
PMID4	B	0.02	0.98
PMID6	B	0.02	0.98
PMID7	B	0.26	0.74
PMID8	A	0.79	0.21
PMID10	B	0.01	0.99

Top Set A

Abstract	Score	Result

Test Set

Abstract	Input	Rank A	Rank B
PMID1	A	0.98	0.02
PMID2	A/B	0.51	0.49
PMID3	B	0.02	0.98
PMID4	B	0.02	0.98
PMID6	B	0.02	0.98

Top Set B

Abstract	Score	Result
PMID10	0.99	B
PMID6	0.98	B