Additional file 1: Performance of all 10 BioReader algorithms and MedlineRanker classifying articles relating to infectious diseases vs. non-infectious diseases (allergy, autoimmunity, cancer, etc.).

|  |  |
| --- | --- |
| **Algorithm** | **AUC** |
| Glmnet | 0.971 |
| kNN | 0.943 |
| SVM | 0.924 |
| SLDA | 0.921 |
| Random Forest | 0.92 |
| Bagging | 0.919 |
| MedlineRanker | 0.912 |
| Max entropy | 0.896 |
| Regression tree | 0.882 |
| Naïve Bayes | 0.879 |
| Boosting | 0.788 |