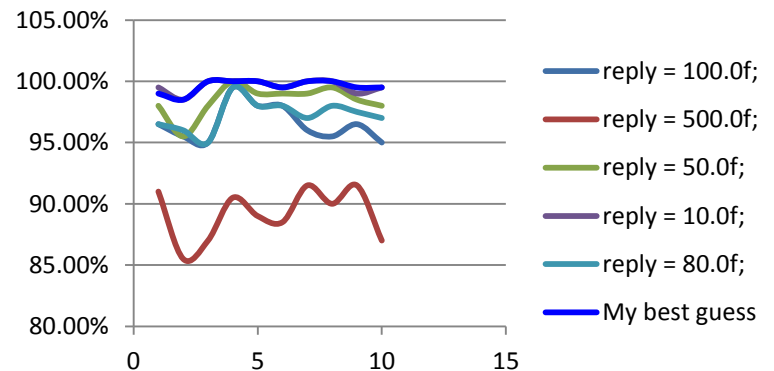
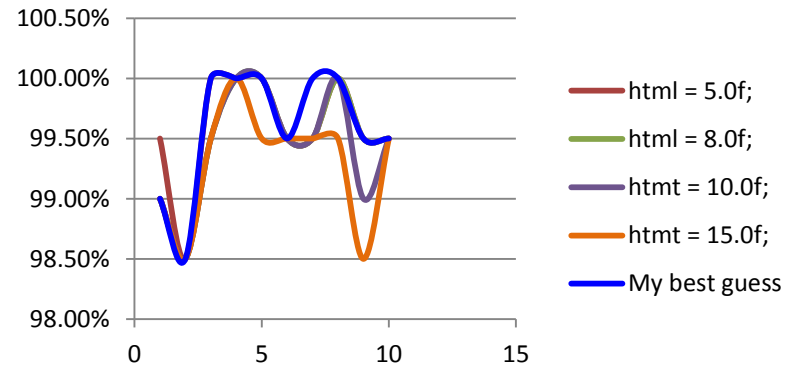
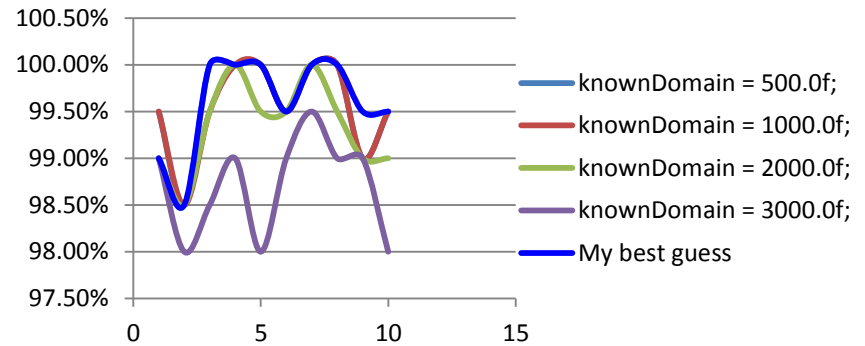
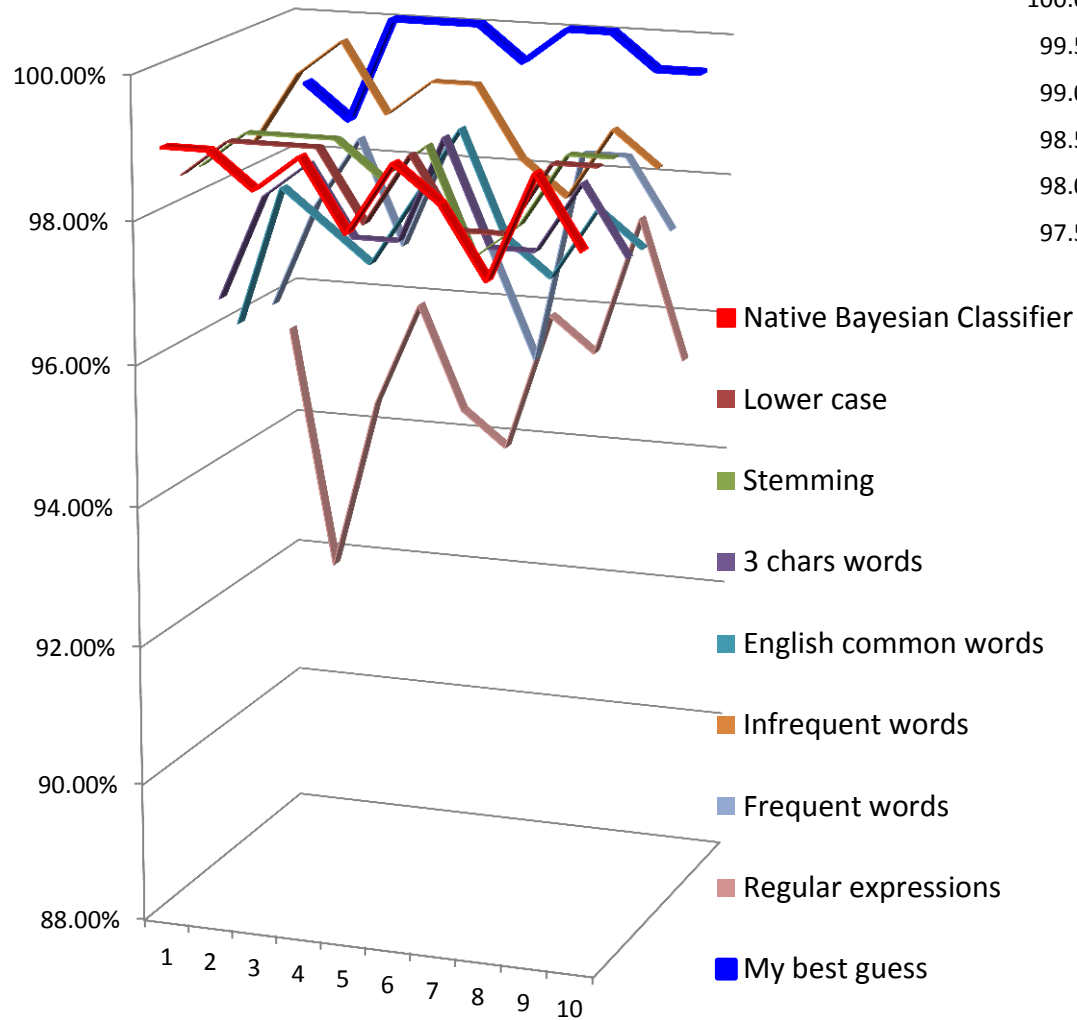
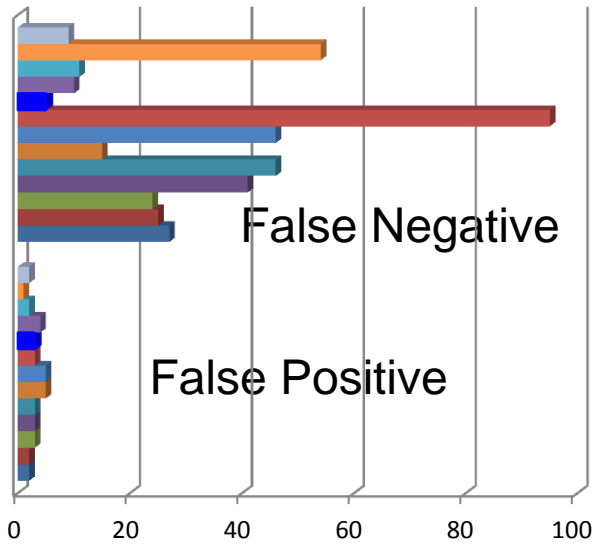




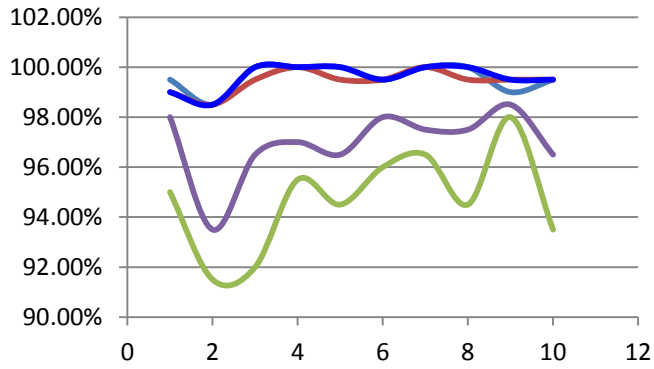
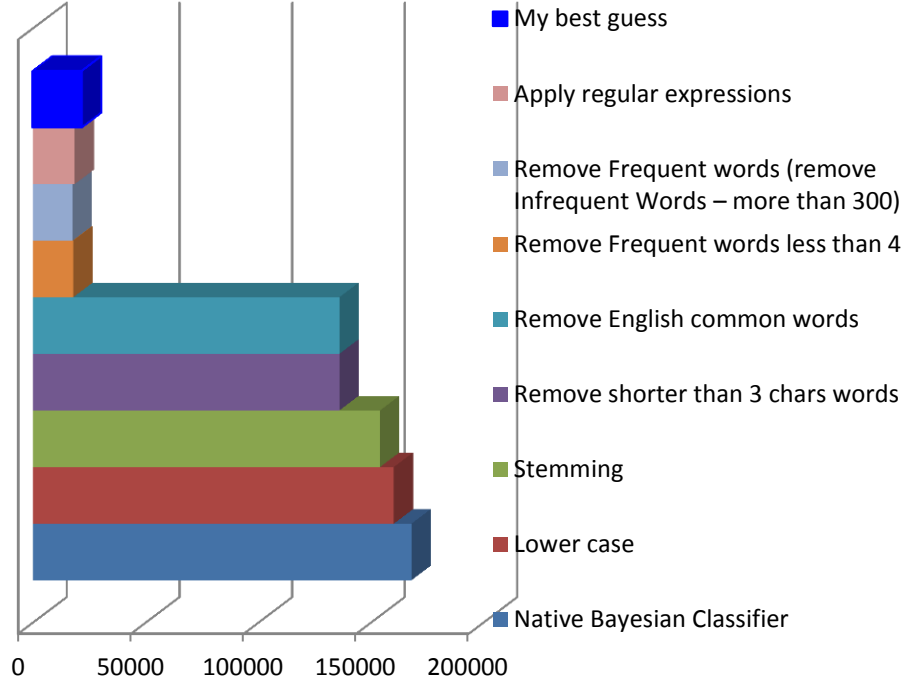
Results from Stratified cross-validation on different techniques





- image = 100.0f;
- reply = 80.0f;
- knownDomain = 2000.0f;
- htmt = 15.0f;
- My best guess
- Regular expressions
- Frequent words
- Infrequent words
- English common words
- 3 chars words
- Stemming

Average Vocabulary Size



T – test showing the difference between Naïve Bayes and the My Best Guess

0.005894292

