

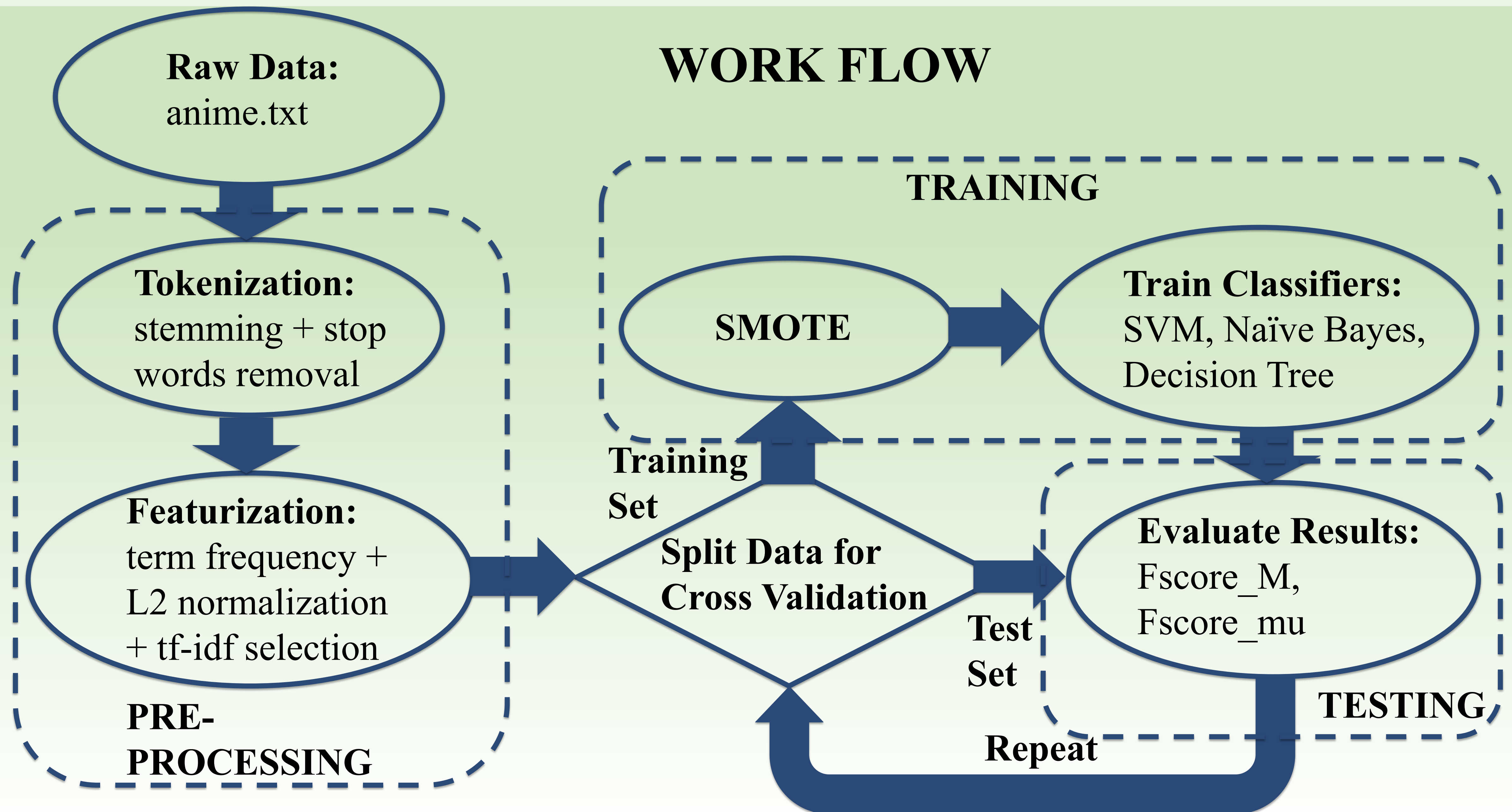
**Do NOT want to add tags for your posts?**

**Need suggestions on your tags?**

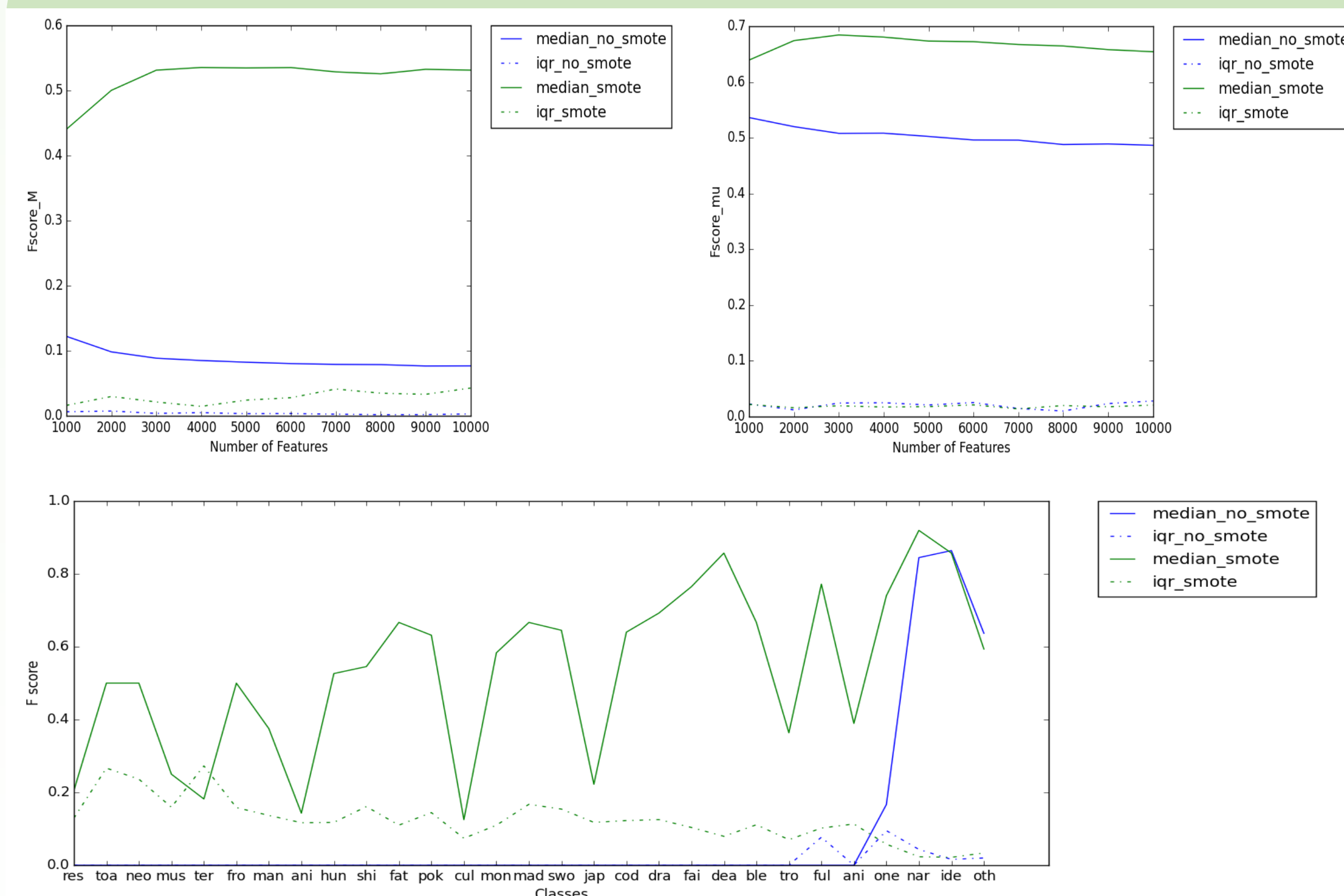
**Look HERE!**

- **Task:** generate tags by prediction
- **Problem:** imbalance
- **Solution:** SMOTE
- **Conclusion:** linear SVM + SMOTE provides the best performance, 5000 features are enough, can provide reliable prediction of tags
- **Future works:**
  - a) SVM with different kernels
  - b) transform multi-classification problem to binary ones and do SMOTE
  - c) test on other data sets

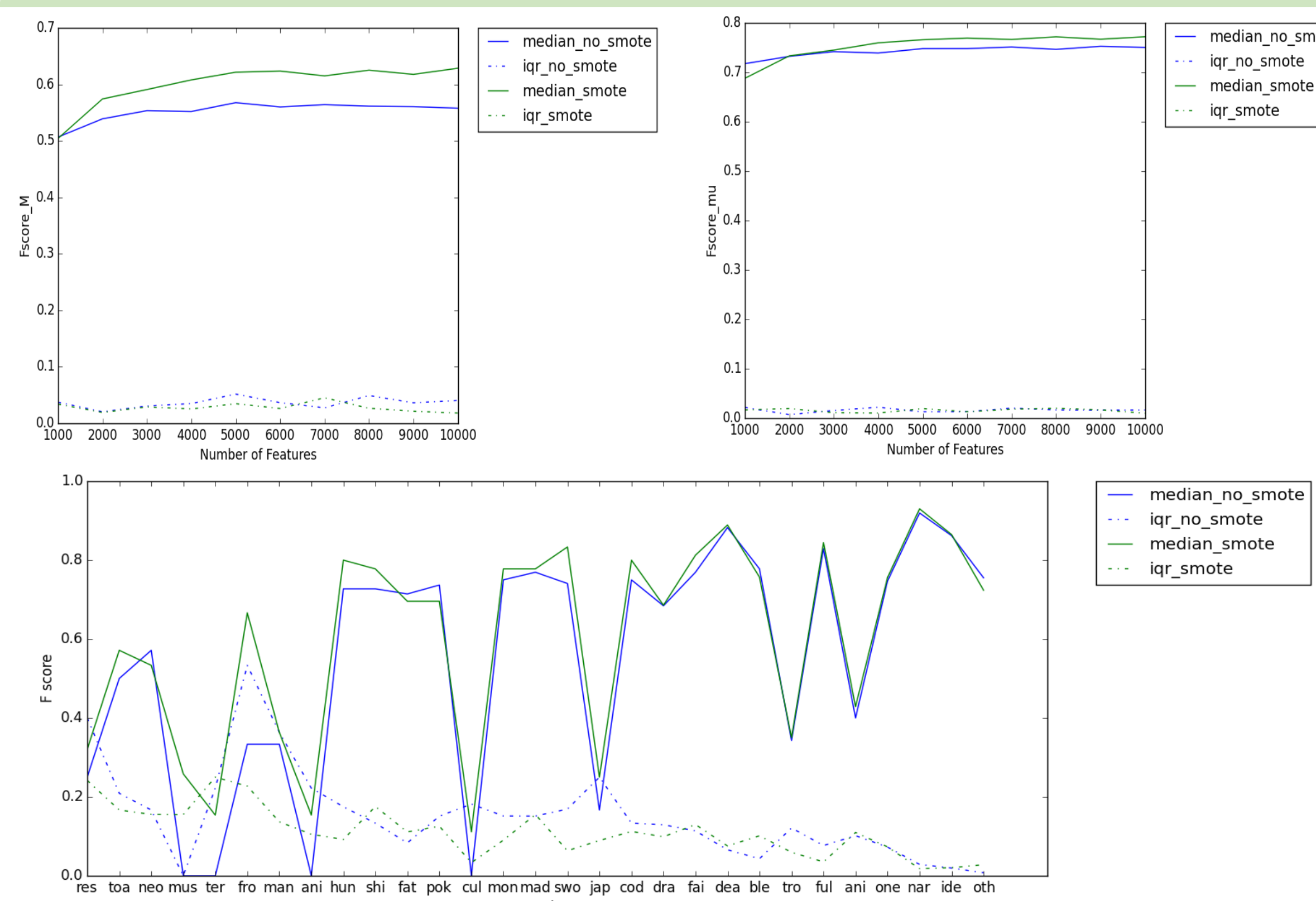
## WORK FLOW



### Naïve Bayes



### SVM



### Decision Tree

