

Progress Summary (Mar 18 - Mar 25, 2020)

Doing

- Write up summary in [Overleaf document](#)

Done

- Read papers: GENERAL, XTREE, SNAP, TimeLIME
- Pre-processing for text mining (not implemented): histogram of most/least used, prune to top-k most frequent words
- Collate defect prediction results using and not using DL from last semester: [Google Sheets link](#)

Todo

- Reproduce Amrit's results from tiny.cc/dodge2020
- Get baselines on text mining data
- Check intrinsic dimensionality of datasets

Roadblocks

N/A