Fall 2017 Progress

Ahmed El-Kishky
Current Project: Subword Mining & Embedding

Details:
- Collaboration with Frank Xu and Aston Zhang.
- Objective is to perform unsupervised word segmentation into sub words, then apply a novel embedding method to combine sub-words into word vector representation.

Progress:
- Implementation of method is complete.
- Writing paper (7 pages written so far) and running experiments.
- Working on submission for Jan. 01 VLDB deadline.
Other Sub-Word Mining Projects

1. **SubWord Mining for Entity Typing.**
   - Main idea is that sub-words provide a strong hint as to the type of the entity in scientific / medical literature.
   - Aiming for a NAACL short paper Jan. 10, hope to extend to a long paper.

2. **Multi-lingual Sub-Word Mining**
   - Evaluating an unsupervised / multi-lingual subword mining algorithm and its applications.
   - Have low-resource languages datasets from Prof. Heng Ji

3. **Supervised Sub-Word Mining**
   - Utilizing small amounts of supervision for subword mining.

   **If you would like to collaborate on some projects, just maybe we can work on something together 😊.**
Awards

I was fortunate enough to be awarded the State Farm Doctoral Fellowship 😊