Study of Classification Uncertainty for Web Service Classification

By Shaun Pinto
Advised by Professor Xumin Liu
Summary

- Need to automate task of web service classification by using machine learning models
- Will help in improve discoverability of API’s thus improving reusability of code
- Models which take into account the uncertainty of the prediction have shown to have high accuracy and have added benefit of reserving judgement
Milestone 3 Goals

- Carry out feature engineering and document how the accuracy is affected by the same

- Build Model that incorporates evidential learning and contrast against other models
Milestone 3 Progress

- Used web crawled information about Source Code, SDK’s etc to further improve the description of the API, (using a similarity measure)
- Compared different embedding algorithms to find out which performed best
- Created a basic model with evidential learning to contrast with other models
# Preliminary Results

<table>
<thead>
<tr>
<th></th>
<th>SVM</th>
<th>Basic Model with Evidential Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original Desc</td>
<td>Mod Desc</td>
</tr>
<tr>
<td>TFIDF</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td>Word2Vec</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>GloVe</td>
<td>0.52</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Conclusions

- GloVe embedding algorithm seemed to perform best
- Improved description seemed to minimally increase accuracy
- Evidential Learning model seem to outperform basic model
- No significant time difference across basic and evidential learning based model
Future Work

- Explore different domains that evidential learning could be used
- Try to improve the model architecture
- Try to carry out more feature engineering / collect better data
Thanks