White paper — Job Skills extraction with LSTM and Word Embeddings

Applying Deep Learning on text corpuses for Job Skills extraction

[](https://medium.com/@nikkisharma536?source=post_page-----d71d1f96024f----------------------)

[Nikita sharma](https://medium.com/@nikkisharma536?source=post_page-----d71d1f96024f----------------------)

Follow

[Sep 7, 2019](https://medium.com/@nikkisharma536/white-paper-job-skills-extraction-with-lstm-and-word-embeddings-d71d1f96024f?source=post_page-----d71d1f96024f----------------------) · 2 min read





Image courtesy — Google

I was recently researching various text mining and language processing techniques to extract Job Skills from Job postings and Resume data. The input data is a free text corpus and the expected output would be the desired skills sets for a given job profile.

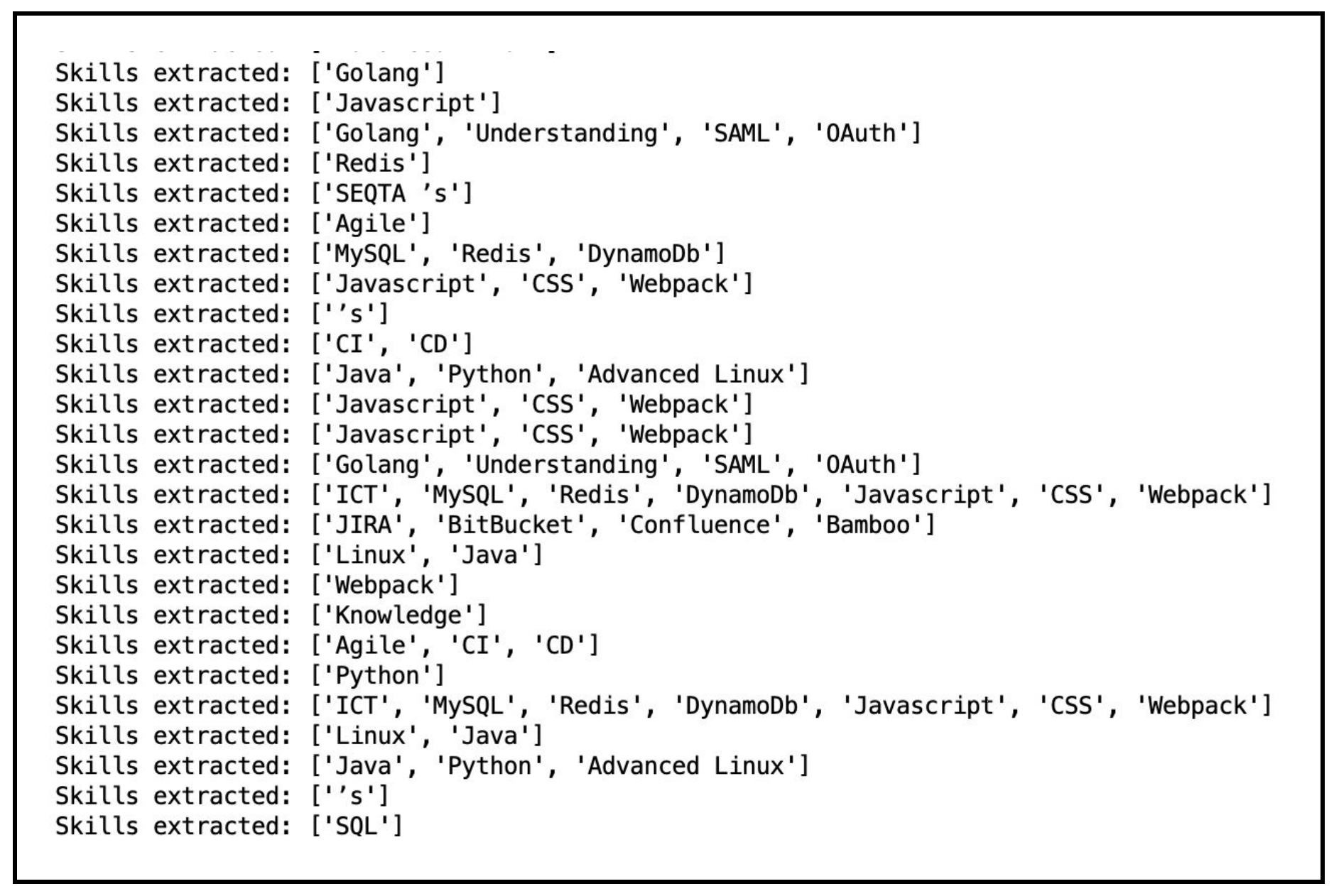
I decided to document all my research as a paper with all the technical details that might be useful for someone researching a similar problem. So here it is -

*Direct link to paper :*[*https://confusedcoders.com/wp-content/uploads/2019/09/Job-Skills-extraction-with-LSTM-and-Word-Embeddings-Nikita-Sharma.pdf*](https://confusedcoders.com/wp-content/uploads/2019/09/Job-Skills-extraction-with-LSTM-and-Word-Embeddings-Nikita-Sharma.pdf)

The output of the exercise were very promising and I was able to extend the model to various Job categories. The techniques is also able to identify new and emerging Skillsets rather than being limited to a known set of Skills.

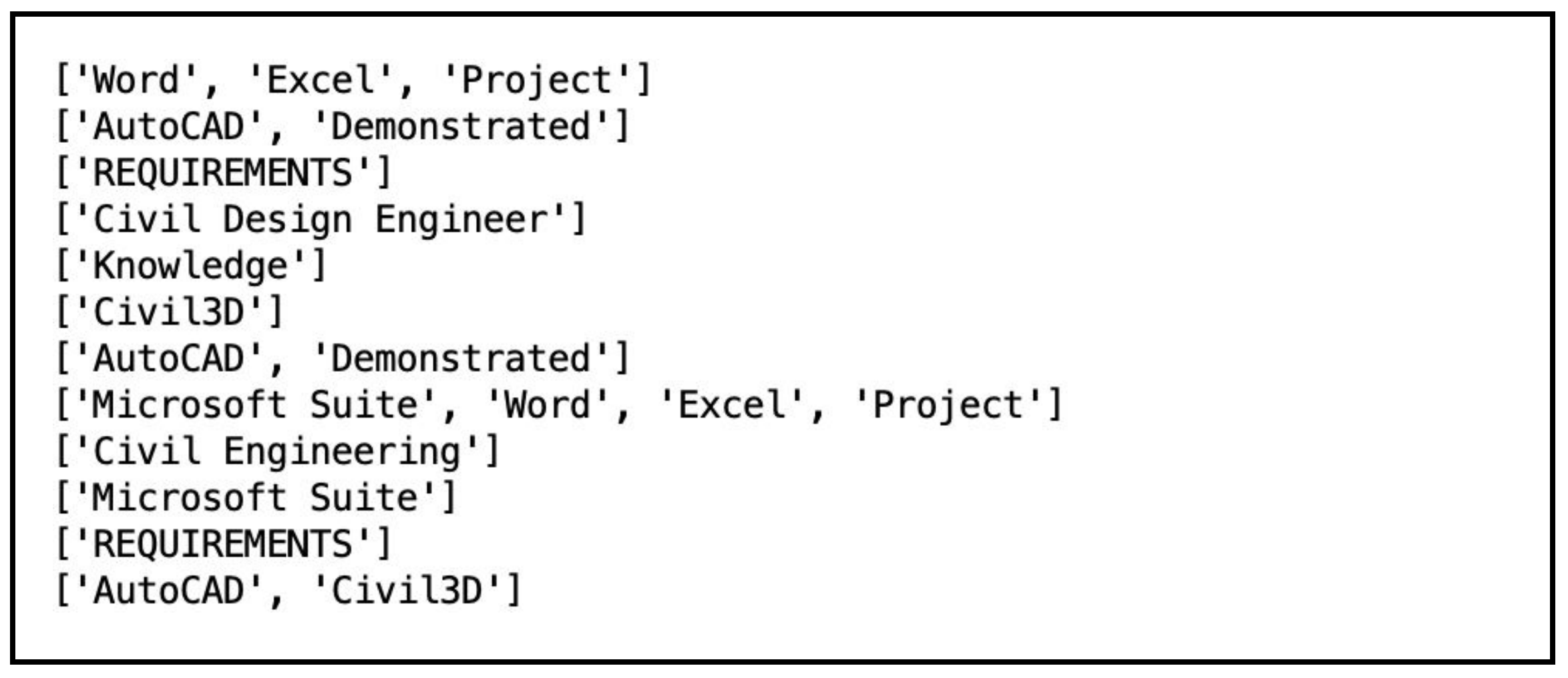
Sample of skills extracted from a Software Engineering Job post:





Same model extended to a Civil Engineering job post:





169

* [Machine Learning](https://medium.com/tag/machine-learning)
* [Deep Learning](https://medium.com/tag/deep-learning)
* [Lstm](https://medium.com/tag/lstm)
* [Word Embeddings](https://medium.com/tag/word-embeddings)
* [Language Processing](https://medium.com/tag/language-processing)