Task 1: Research Topic + Three References

Write a forum post below, explaining the problem or question that you want to focus on, and how you are planning to solve that problem or answer the question. You don't need to be specific - the only important thing is that everyone should have a somewhat suitable topic in mind for next week. Please name your post by the preliminary title for your study!
Also try to find a few references that you could use in your study. List those in your submission.
You can think about the project through some of the questions in `what?' and `why?' quadrants of the `Watson box':

What?	*Why?*
What am I really interested about? What do I want to know better or understand better? What questions do I think are really important to be solved?	Why is this going to be interesting or important to others? Who will be interested in reading about what I found? Does this reveal something that we didn't previously know? Why am I the right person to do this?

Learning objective: To get everyone started with their research projects.
Min. number of words: 150

Word-Embeddings: Encoding Algorithms and Evaluation Techniques

Word-embeddings are language modeling techniques representing natural language words or phrases as real-valued vectors. What is special about these vectors is their semantic interpretability: words with similar meanings get close or computationally related values. This allows performing arithmetic operations on the encoded vectors to guess words following a certain relational pattern. For example, the question 'What is to France what Tokyo is to Japan?' can be answered with the following linear equation: Japan – Tokyo + France = X, where the value of X is mapped to the word it is closest to from a set of possible target answers, which in this case should be 'Paris'.
In this research, we aim to review the state-of-the-art algorithms and approaches used in generating and evaluating word-embeddings. We also hope to review the major applications of this technique and its current limitations. This topic is important because of the seminal role word-embeddings played in NLP research over recent years, particularly in enabling the introduction of deep learning techniques to the field. Our team is well placed to explore this topic because of our previous exposure to ML techniques (particularly deep neural networks) and our interest in further exploring such fields in future courses and research work.

References:

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781. https://arxiv.org/abs/1301.3781
Pennington, J., Socher, R., & Manning, C. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543). https://www.aclweb.org/anthology/D14-1162/
Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics, 5, 135-146. https://www.mitpressjournals.org/doi/pdfplus/10.1162/tacl_a_00051
Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759. https://arxiv.org/pdf/1607.01759.pdf