**NLP resources:**

Best way to learn and quickly get results in hackathon is by ‘doing’ (i.e. using tools already written). Here are some below…

Some papers that may be useful:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5001761/>

<https://www.ncbi.nlm.nih.gov/pubmed/28830409>

How to cluster sentences for patterns:

<https://eight2late.wordpress.com/2015/07/22/a-gentle-introduction-to-cluster-analysis-using-r/>

<http://science.slc.edu/~jmarshall/courses/2005/fall/cs151/lectures/RAAM/>

Mapping to UMLS, NCIt, SNOMED, RxNorm,etc:

<http://noble-tools.dbmi.pitt.edu/>

<https://metamap.nlm.nih.gov/>

Online forum comment:

“Since your input is in the natural language form, best way to start looking into it, first by parsing the sentence structure. and running the sentence through NER (Named Entity Recognizer).

Parsing the sentence lets you come up with rules such as, certain types of dependencies always give you the intent. Running the NER will let you identify places and dates. If it's not simple to come up with rules to classify the intent, you can as well use a classifier to do the same using feature vector formulated from the input sentence. In fact some of the parser out put can go into formulating the feature vector.

For both there exists software's from [Stanford NLP Group](http://nlp.stanford.edu/)

May be you can look into:

* [Stanford parser](http://nlp.stanford.edu/software/lex-parser.shtml)
* [Stanford NER Tagger](http://nlp.stanford.edu/software/CRF-NER.shtml)

Once you parse the sentence, you have intent and other information require to answer the question.

Ex: I took your sentence "Will it be sunny this weekend in Chicago." and ran it through [Online Stanford NER Tagger](http://nlp.stanford.edu:8080/ner/process). Which gave me the following:

Will it be sunny this <DATE>weekend</DATE> in <LOCATION>Chicago</LOCATION>

Now you have identified date and location.

I hope this helps. I know the answer is quite generic, and may be helpful in just getting started.”

See scikit learn and google distance:

<https://www.google.com/amp/s/amp.reddit.com/r/MachineLearning/comments/2qkjd5/how_to_cluster_text_sentences_unsupervised/>

<https://stackoverflow.com/questions/47818489/how-to-use-clustering-to-group-sentences-with-similar-intents>



<https://github.com/RxNLP/PyRXNLP>

<http://www.nltk.org/book/>

Clustering for templates (example paper):

<http://www.surdeanu.info/mihai/teaching/ista555-fall13/readings/chambers2011.pdf>

<https://nlp.stanford.edu/software/>

Nlp background:

<https://blog.insightdatascience.com/how-to-solve-90-of-nlp-problems-a-step-by-step-guide-fda605278e4e?gi=87f35fc58253>