

Heyday x UBC Capstone project

 heyday
by Hootsuite®



Timeline

- May 3, 4: Group agreement & Hackathon (UBC)
- May 5: the official first day of capstone - Heyday x UBC
- May 5 - June 15: Capstone project
 - May 5-11: Group agreement & project planning & research
 - May 12-June 15: Working on the project
 - Every Thursday afternoon (EST), a feedback & Q&A session - Heyday x UBC
- June 15 - June 29: Project wrap up & students will prepare for the final report and presentation
 - Recorded presentation or live presentation @ Heyday's KS session [TBD]



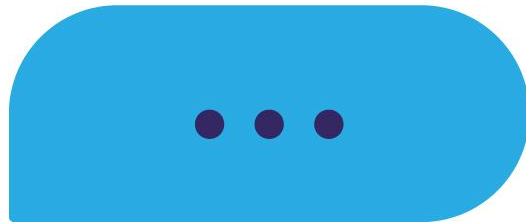
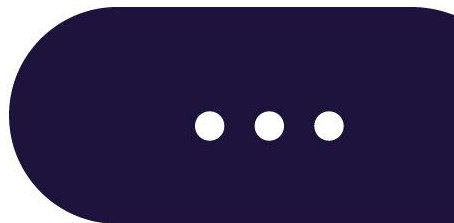
Expectation

- Developing a negation-aware named entity recognition system in the **product search** context
 - Thus, the data should be product search related utterances
 - Our past approach: [neg-spacy](#) library
 - Limitation: doesn't work well with our domain, doesn't work well with relatively longer sentences
 - The validation score with this library can be a baseline score
 - We can try beating this baseline score through this project!
 - NER labels of our interest:
 - Size, Colour, Category, Quantity, Brand, etc [TBA]



Expectation (cont.)

- Some other stuff:
 - We will provide a validation set , and there will be a hidden test set ;)
 - Final score with a hidden test set will be revealed at the end of the project
 - The project will mainly focus on:
 - Creating training dataset & exploring public datasets
 - Modeling (training, fine-tuning)
 - Creating evaluation framework





Data structure

```
{
  "text": "I want to buy a red hat without the blue ribbon",
  "entities": [
    {
      "text": "red",
      "entity": "colour",
      "start": 16,
      "end": 19,
      "isNegative": False
    },
    {
      "text": "hat",
      "entity": "product",
      "start": 20,
      "end": 23,
      "isNegative": False
    },
    {
      "text": "blue",
      "entity": "colour",
      "start": 36,
      "end": 40,
      "isNegative": True
    },
    {
      "text": "ribbon",
      "entity": "product",
      "start": 42,
      "end": 48,
      "isNegative": True
    }
  ]
}
```

