



# NLP Tutorial Project

## MATH 497

Alexis GAZIELLO  
Christopher WEICONG HONG  
Rahul MARAMREDDY

Instructor: Robert Ellis  
Tutorial Author: Emily Webber

# Index

— — —

1. Problem Statement
2. Data Overview
3. Custom model build
4. Topic Modeling
5. Amazon Comprehend
6. Conclusions

# Problem Statement

---

- Millions of Amazon reviews that product owners want to keep track of
- Classify reviews based on sentiment (Logistic / LSTM / AWS Comprehend)
- Understand common trends of negative reviews (Latent Dirichlet Allocation)

**Benefit:** Reusable code for text classification, topic modeling, and AWS Comprehend

# Data Overview

Amazon product review

Label	Title	Content
Negative	This book was sooo bad :(	I didn't enjoy reading it because (...)
Positive	(...)	(...)
(...)	(...)	(...)
Negative	(...)	(...)

Source: <https://course.fast.ai/datasets>

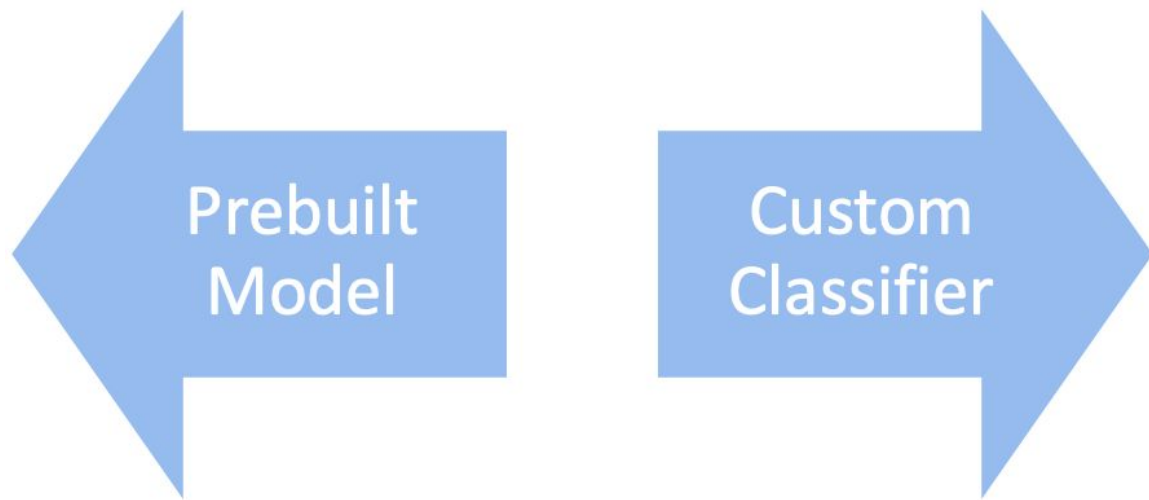
(Amazon Reviews: Polarity)

# CUSTOM MODEL

# TOPIC MODELING

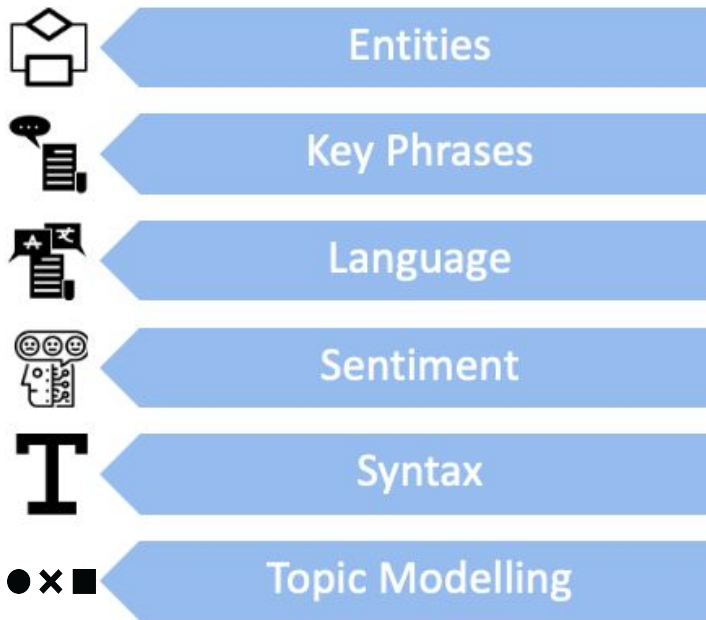
# Amazon Comprehend

— — —



# Amazon Comprehend Pre-built Model

— — —





# AMAZON COMPREHEND

# Conclusion

---

	Comprehend Prebuilt Classifier	Comprehend Custom Classifier	Custom Model
Description	No assembly required	Little assembly required	Build on your own
Knowledge Required	No knowledge necessary	Need of basic DS knowledge (preprocessing)	Need of a LOT of DS knowledge
When to use it?	Quick solution for unlabeled data	Quick solution for labeled data	Complicated solution for any situation
Customization	Very limited	Limited customization	Unlimited
Price	Really low	Low	Expensive
Target	Small-medium sized companies	Small-medium sized companies	Large companies

Questions?