Build a Customized NLP Service

Wednesday, June 2, 2021

Presented by: Yongchao Zhou

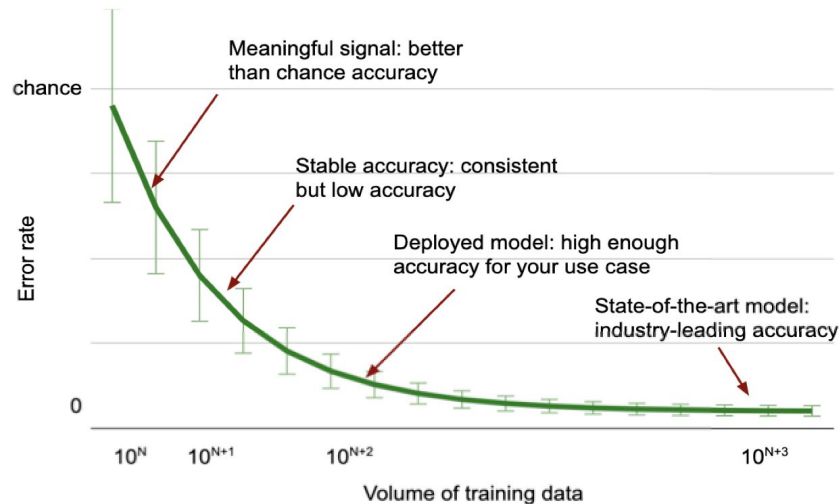VECTOR INSTITUTE

Background

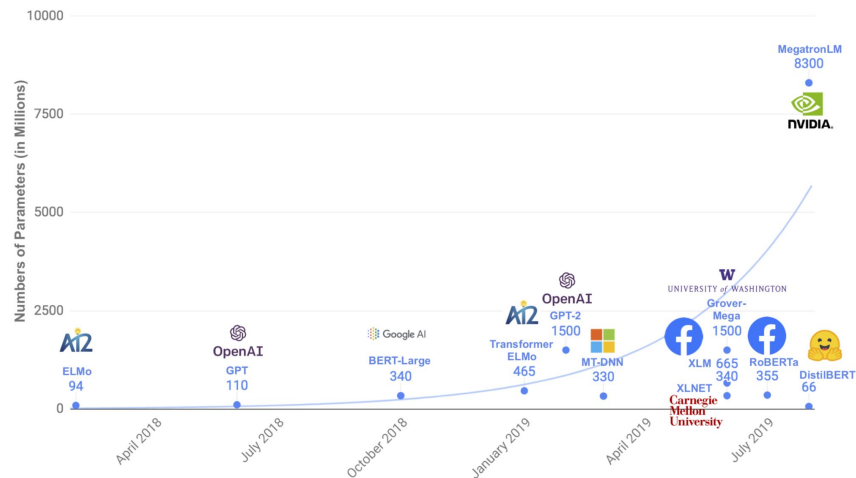
Figure 1: Error Rate vs Volume of Training Data

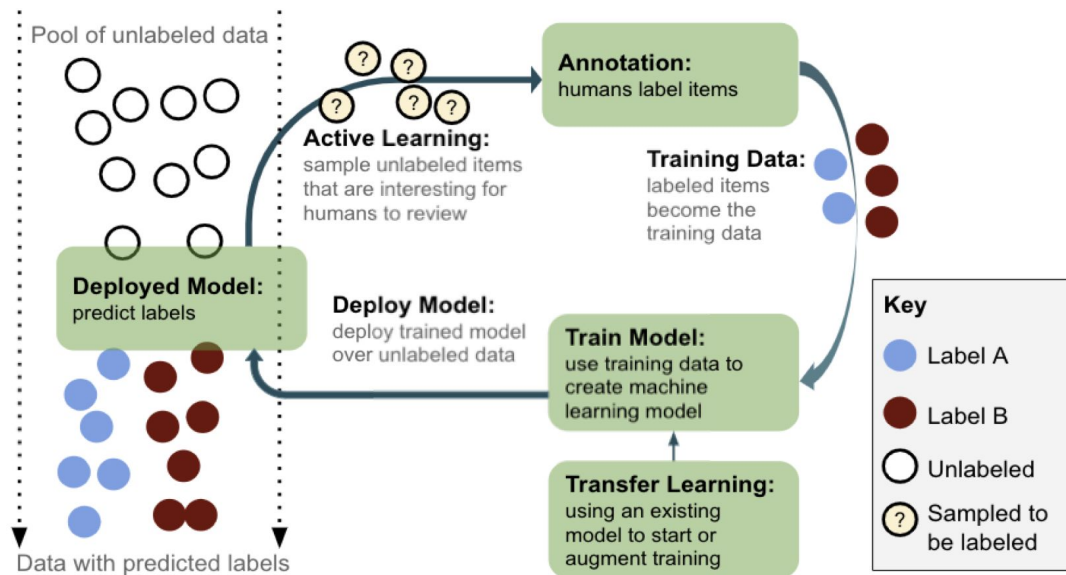Figure 2: Machine Learning Model Size

Background

Figure 3: Active Machine Learning + Transfer Learning Workflow

Active Learning (AL)

+

Transfer Learning

Data Annotation Tool

+

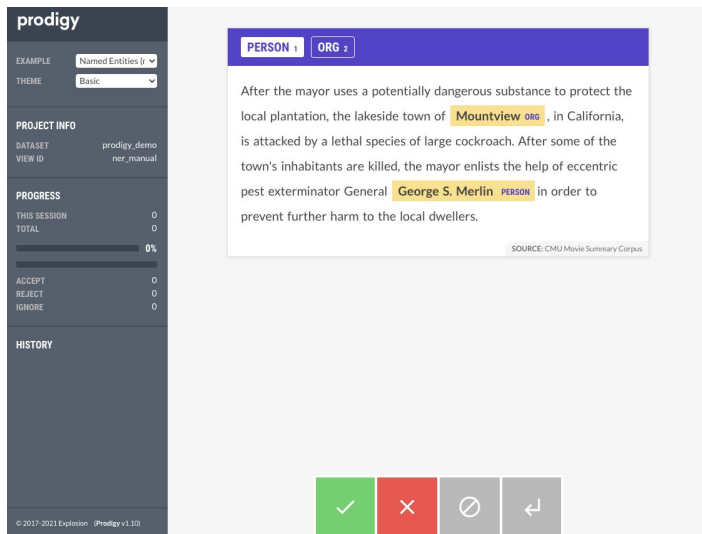Vector's NLP Services

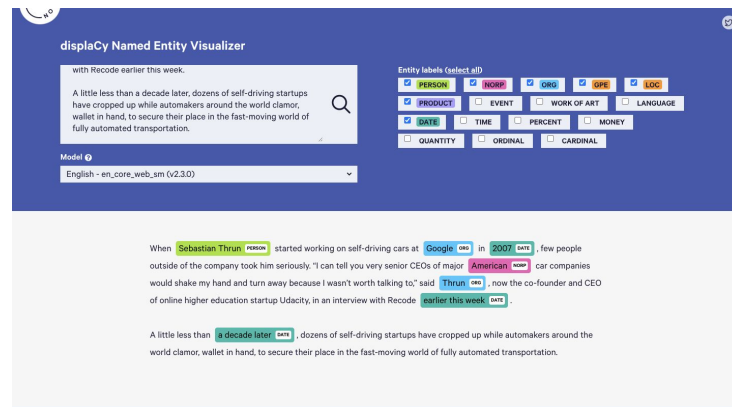# Example: Customized NER Service

Figure 4: Annotation Tool

Figure 5: NER Service

https://explosion.ai/demos/displacy-ent
https://prodi.gy/demo

VECTOR
INSTITUTE

What are missing?

- Data Annotation Tool
  - Most tools do not use AL and Good Tools are not open source
  - No control over the underlying AL algorithm
    - Advanced AL algorithm with Uncertainty and Diversity Trade-off
    - Auto-Suggestion Strength
    - User Preferences
- NER Service
  - Customized Model + Customized Entity Support (Law, Medical, Finance)

Deliverables

- NER Annotation Tool
- Customized NER Service Interface

- A Guideline/Tutorial/Demo on how to build a customized NER service based on active learning and transfer learning

Thank you!

VECTOR INSTITUTE