Techniche

Lee Mackey

Techniche- general summary

Decision support tool using patent texts to predict technological specializations - “tech niches” - in machine learning markets

Tech niche?

- Specific sub-category or field in ML technology/market

- Specialization in terms of multiple types of entities
  - company (assignee)
  - location
  - inventor

Business value and users

- **Value**: Business decision-makers need knowledge of pipeline of machine learning technologies that might shape markets into the future

- **Granularity:** finding sub-categories below patent level

Users

Business consumers in ML space who want to evaluate probability and significance of future market trends to inform management decisions

- analysts conducting competitive analysis
- investors valuing companies via IP assets
- attorneys preparing IP cases
- policymakers supporting local business clusters

Data

- **Data**: Text data and metadata of public patent documents from US Patent Organization (USPTO) - PatentsView API
- **Preprocessing**: Basic preprocessing techniques
- **Data for specific task**: 1000 patents

Initial multi-class classification model

- **Task**: given text data from abstract and title of patent invention in natural language processing (NLP) space, predict company (assignee) that owns patent
- **Features**: abstract + title text **Target**: patent assignee
- **Result**: no better accuracy than baseline of predicting most frequent class, across 13 classes

Next steps

- Data size increase
- Compare:
  - Word representations/embedding
  - Hyperparameter settings
- Predict labelled technological categories
- Compatibility across cpu/gpu, vm images, local/VM
- Additional tasks