

PCATx CORE

A web crawling and artificial intelligence framework for Praedicat, Inc. developed by the 2018 RIPS team at UCLA's Institute for Pure and Applied Mathematics. Our goal utilize Natural Language Processing techniques and Artificial Intelligence to collect automate data collection for the InsurTech company.



Table of Contents

- Classifier
- KPM == Keyword Producing Model
- Parser
- Query Formulator

Classifier

Mapper and Keyword Producing Model (KPM)

This module is divided into two parts: the Mapper and the Keyword Producing Model (KPM).

Mapper is essentially a wrapper for a dictionary, it associates unique identifiers with a list of aliases for the entity identified. An example would be "00001" -> ["Micron Technology", "MU", "Micron"]. For full documentation on Mapper, go to [docs/Mapper.md](#).

The Keyword Producing Model takes data and finds words highly associated with our entities to build the list of aliases using data2vec and neural networks.

The following modules are part of the Keyword Producing Model:

- [corpusBuilder](#) --- [Documentation](#)
- [Mapper](#) --- [Documentation](#)

Parser

Query Formulator

