

Formal Specifications from Natural Language Requirements

ABSTRACT

Ambiguity is often observed in technical specifications because they are written in a natural language (e.g., English). This project aims to develop an automated approach to formulate formal specifications from natural language requirements, such as regulation and policy documentation.

GOALS

- Develop an automated approach to formulate formal specifications from natural language requirements (e.g., regulations, policies).
- Gain a better understanding of Natural Language Processing (NLP) and how it works to eliminate ambiguity in technical documentation.

TECHNOLOGIES



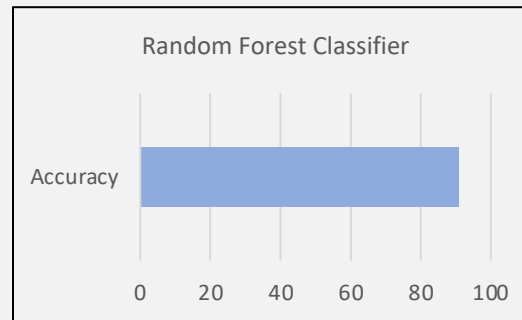
Python



Stanford Typed
Dependency Parser

FINDINGS

As a group, we manually classified several sentences as temporal or non-temporal. We then utilized a Random Forest classifier to determine accuracy, which was 90.9%.



IMPACT

- Contribution of knowledge will serve to advance the field of NLP.
- Automating the classification of requirements will reduce ambiguity in engineering projects.
- Bridging the gap between natural language and formal specification will increase productivity in today's Agile approach to software development.
- Proving ties between temporal phrases and requirements will aid Project Managers and Developers to more accurately estimate their time.

CHALLENGES

- Each group member is relatively new to the field of NLP. Limited knowledge caused time constraints during early stages.
- Comprehension of highly technical papers and knowing what information to extract.
- Deciding on Python libraries for our machine learning algorithm.

ACHIEVEMENTS

- Learned how to use variety of NLP tools.
- Implemented a protocol to determine temporal phrases in requirement documentation.
- Confirmed correlation between our findings and those presented by researchers at the Jet Propulsion Laboratory.



Dr. Nan Niu,
Advisor



Garret Oien,
Member



Kevin Starmack,
Member



Sheffi Tiwari,
Member