

Similarity Predictor

October 01, 2017

Overview

Devise a system that can calculate similarity score between different text chunks. This similarity will also take into consideration contextual overlap apart from text based similarity.

Process Flow

- A web interface that will have two text areas to capture two different text chunks.
- Once done user will hit Submit button
- Data will be received by the backend server which will do all the operational logic and return results to the web interface
- Web interface will display a percentage score between 1-100%

Data Source

- You can use data from <https://patents.google.com> labelled under **Claims** header
- For benchmarking pick any two claims from same Patent. They will usually have high similarity as they are defining same concept most of the time.
Like Claim 1 and Claim 11 of [US8708211B2](#)

Claim 1

A powered surgical apparatus for engaging tissue, comprising:

a handle assembly having a first motor and proximal and distal portions;

a movable portion operatively connected to the distal portion of the handle assembly, the movable portion being movable with respect to the handle assembly;

a tool assembly adapted to engage tissue, the tool assembly being operatively coupled to the movable portion;

a power source configured to supply electrical power; and

a transmission system configured to transfer electrical power between the handle assembly and the movable portion, the transmission system including a first electronic board disposed in the handle assembly and a second electronic board positioned in the movable portion, the second electronic board including a primary control circuit and a secondary control circuit, wherein the transmission system transmits control signals between the first and second electronic boards, the first electronic board controls the first motor in response to a control signal received from the second electronic board, wherein the movable portion is configured and adapted to rotate about a longitudinal axis defined by the powered surgical apparatus.

Claim 11

A powered surgical apparatus for engaging tissue, comprising:

a handle assembly having a first motor;

a movable portion operatively connected to a distal portion of the handle assembly, the movable portion configured and adapted to rotate about a longitudinal axis defined by the powered surgical apparatus, the movable portion has a second motor;

a tool assembly operatively coupled to the movable portion; and

a transmission system including a first electronic board disposed in the handle assembly and a second electronic board positioned in the movable portion, the second electronic board including a primary control circuit and a secondary control circuit, wherein the transmission system transmits control signals between the first and second electronic boards.

Milestones

Stage 1

- Discuss the strategy with GreyB after conducting a thorough research. Also shortlist programming stack

Stage 2

- Divide project in small subtasks along with the delivery dates
- Trial and test if you have shortlisted multiple strategies
- Share results with GreyB

Stage 3

- Start implementing a simple web interface as described in the specifications