Automated Text Tagging and Summarization

Introduction

- Filtering and Tagging of Text Data
- Automated Text Summarization

The Problem

- Content gets generated constantly
- Internet users look for keywords and skim text
- Search engines respond better to pages with tags

The applications of better natural language processing tools are endless

Who is this for ?

- Client - Local online newspaper platforms
- Customer - Editorial staff
- User - Publishers

Our Solution

Our system consists of a web frontend and API based on a python library.

Extractive Techniques

- TF/IDF
- Sentence Ranking and Elimination

Abstractive Techniques:

- Word Embedding
- Sentence Reduction

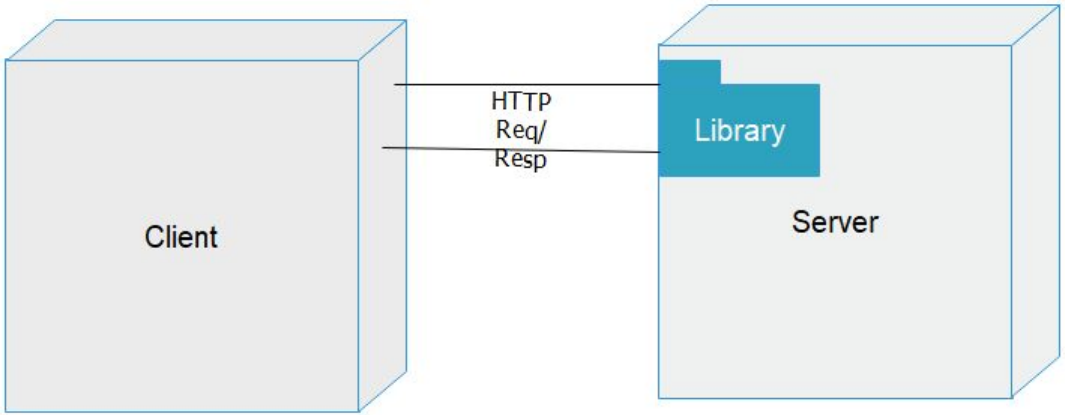
The Algorithm(s)

- Sentence Boundary Disambiguation
- Stop Word Removal
- Word Stemming
- Sentence Reduction
- Text Summarization
- Word Embedding

Benefits

- Increased website visibility
- Increased traffic
- Decreased time-to-delivery of content

Design



Implementation

- Frontend
- Backend
- Library

Development approach

Agile

- Sprint 1
- Sprint 2
- Sprint 3

Testing

- Formal specification and verification
- Functional
- Load testing

Maintainability

- Client/Server Architecture

Demo

Questions ?