



Automatic Keyword and Sentence-Based Text Summarization for Software Bug Reports

21266013	Moumita Khandaker
21266015	Asadullah Al Galib
22166040	Moh. Shohanur Rahman
22166049	Sayed Bin Amirul Sakin
22166053	Mofiz Mojib Haider
1034371479	Zaber Mohammad



Introduction & Motivation

- Unsupervised approach to summarize bug reports
- Rapid Automatic Keyword extraction and term frequency-inverse document frequency method
- Fuzzy C-means Clustering
- Reduction of time and effort of software testers



PRELIMINARY CONCEPT

Bug Report Summarization Concepts



TEXT PRE- PROCESSING


- Segmentation
- Tokenization
- Removal of stop words
- Removal of Punctuation



DTM CONVERSION

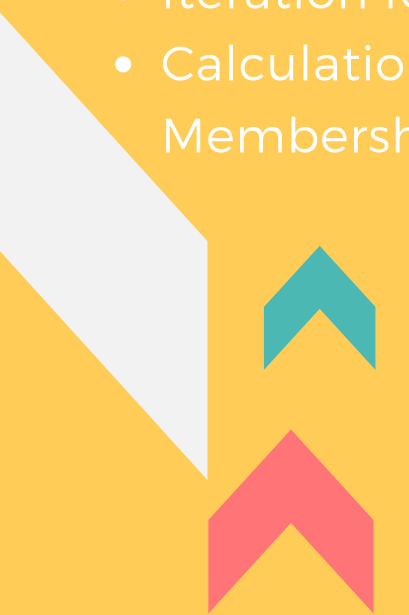


RAPID AUTOMATIC KEYWORD EXTRACTION

- Split into word array
 - Split of array into contiguous words
 - Score assignment to candidate keywords
- 



Fuzzy Clustering

- Application of Fuzzy C-means Clustering
 - Cluster Center Selection
 - Assignment of membership value
 - Iteration for Cluster Center Updates
 - Calculation of new Degree of Membership
- 

Hierarchical Clustering

- Separate Cluster Consideration of each data point
- Proximity Distance Calculation
- Merging of similar proximity cluster
- Iteration

AVERAGE LINK METHOD

- **PROXIMITY CALCULATION
BETWEEN CLUSTERS**
- **AVERAGE DISTANCE
CALCULATION**

Research Methodology



- Corpus Creation and Text Pre-Processing
- Feature Extraction
 - i. Keyword Features
 - ii. Sentence Level Features
- Rule Engine
- Hierarchical Clustering





Implementation and Results

- Data Collection
- Evaluation Metrics
- Process Illustration



Threats to Validity



The diagram consists of two large, light gray, arrow-shaped boxes pointing to the right, stacked vertically. The top box is labeled 'Internal Threats' and the bottom box is labeled 'External Threats'. Both boxes are set against a solid red background. To the left of the top box is a yellow arrow pointing right. To the right of the top box is a teal arrow pointing right. Below the top box and above the bottom box is a white arrow pointing right. To the left of the bottom box is a yellow arrow pointing right. To the right of the bottom box is a teal arrow pointing right, and further to the right is a white arrow pointing right.

Internal Threats

External Threats



”

Related Work and Conclusion

- 
- Extractive Summarization
 - Bug Summarization
 - Term frequency-inverse document
 - Rapid automatic keyword extraction

“

