

Workshop 2

Kaung Khant Lin

6540131

541

Task 2

Q1

- Stopword removal changed the text the most

Q2

- Search Engines: Stemming
 - Faster and finds more matches
 - "running", "runs", "runner" all become "run" → better search results
- Machine Translation: Lemmatization
 - Keeps real words that can be translated properly
 - "memories" → "memory" (real word) vs "memori" (not a word)
- Sentiment Analysis: Lemmatization
 - Preserves meaning better for understanding emotions
 - "disappeared" keeps its full meaning vs "disappear" which loses context

Task 3

1. $(0|1)((0|1)(0|1))^*$
2. $[a-z]^*b$
3. $^(b(b|ab)^*)? \$$