

11/21/2024

Analysis of Repetition in Teaching Summarize Findings and Steps,

Analyzing Similarity Between Two Courses

Objective:

This prototype is designed to assess the content similarity between two different courses (focusing on AI and ML topics). The purpose of this analysis is to identify overlaps and repetitions in these courses to improve content and reduce redundancies.

Steps Taken:

1. Course Selection:

Two text files containing lecture content from two courses were selected for analysis (I downloaded teams)

- **Course 1:** *AI2.docx*

- **Course 2:** *Vir1.docx*

2. Defining Specialized Keywords:

A set of domain-specific keywords related to AI and VIR was defined for analysis:

- **AI:** 'ai', 'machine', 'learning', 'intelligence', 'algorithm', 'data', 'model'

- **VR:** 'story', 'narrative', 'data', 'visualization', 'insights', 'message',

3. Text Preprocessing and Cleaning:

Both files were processed to remove noise and non-useful patterns, ensuring a standardized and analyzable format.

4. Keyword Analysis:

The frequency of keywords related to AI and VIR was calculated in each text:

- **AI2.docx:** 82 keywords related to AI and 27 keywords related to ML

- **Vir1.docx:** 123 keywords related to AI and 216 keywords related to ML

5. Similarity Calculation:

Using basic formulas, the similarity score between the two courses based on keywords was

computed: (without using Keyword Analysis I have Similarity Score : 0.70 so I decided to use Keyword based on search on the google)

◦ **Similarity Score:** 0.2433

6. **Word Cloud Generation:**

A Word Cloud was generated for each course, visually displaying the distribution of keywords within the texts.

Findings:

This prototype indicates a meaningful similarity between the two analyzed courses (0.2433). The shared frequency of keywords suggests the possibility of overlapping or repetition in the teaching of certain concepts.

***I think we should use LMS because this data has a lot of irrelative words**