## Overview

This application is a Virtual Threaded Text Simplifier that processes text files using advanced Java threading and data structures.

It leverages modern Java features, such as virtual threads and concurrency, to handle large datasets efficiently.

## Main Features

---

**Text Simplification:**

The application simplifies input text files by replacing words with their most semantically similar counterparts using pre-trained word embeddings.

The similarity is calculated using cosine similarity, ensuring accurate replacements within a specified threshold.

**Word Embeddings Integration:**

Reads a large embeddings.txt file containing word vectors to map words to numerical representations.

The similarity threshold can be customized to fine-tune the replacements, balancing precision and flexibility.

**Google 1000 Word Compatibility:**

The application integrates with a curated google-1000.txt file for matching and simplifying common words.

It outputs enriched files where matches are updated while preserving the original formatting (e.g., punctuation and newlines).

**Virtual Threading for Efficiency:**

Processes files concurrently using Java virtual threads, enabling faster execution on multi-core systems.

Handles large datasets without overwhelming system resources, making the application scalable for real-world use.

**File Handling:**

Supports user-specified input, output, embeddings, and reference files.

Ensures the integrity of output files by appending changes without overwriting existing content.

**Customization & User Interaction:**

A command-line menu allows users to customize file paths, adjust similarity thresholds, and execute tasks interactively.

Offers optional debugging outputs for advanced analysis.

**Code Modularity:**

Features reusable components like EmbeddingUtils, Simplifier, and WordProcessor for maintainability and scalability.

**References:**

*https://docs.oracle.com/en/java/*

*https://adoptium.net/*

*https://stackoverflow.com/*

*https://mvnrepository.com/*

*https://www.baeldung.com/*

**AI Assistance Prompt (Chat GPT):**

*"What is the difference between a virtual thread and a platform thread in Java?"*

*"When should I use a HashMap vs a ConcurrentHashMap?"*

*"What threshold for cosine similarity ensures realistic word replacements?"*