

Prerequisites

Before running the project, ensure you have the following installed:

1. **Java Development Kit (JDK):** Version 8 or higher.
2. **Visual Studio Code (VSCode):** A lightweight but powerful source code editor.
3. **Java Extension Pack for VSCode:** Install it from the Extensions Marketplace in VSCode.

Setup Instructions

1. Download the Zip file and Unzip it

- Unzip the zip file in your system

2. Open the Project in VSCode

1. Open **Visual Studio Code**.
2. Go to **File > Open Folder** and select the folder containing the project files.

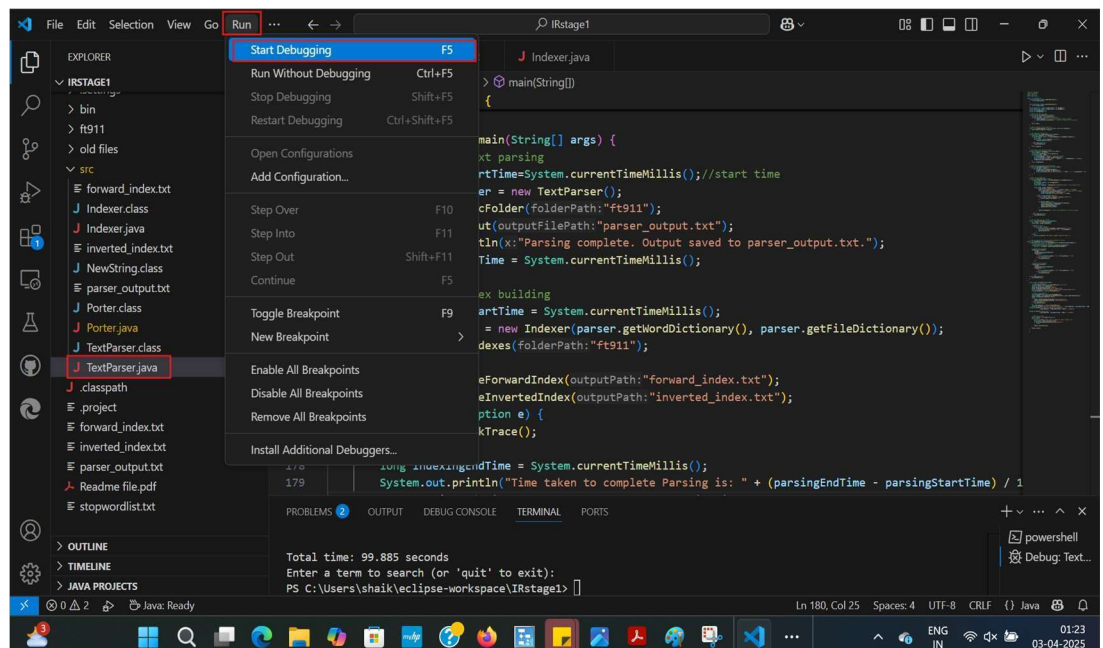
Running the Project

1. Compile the Java Files

1. Open the **Integrated Terminal** in VSCode:
 - Go to **Terminal > New Terminal** (or press `Ctrl + ``).
 - Navigate to your respective src folder. In my case it is: **cd C:\Users\shaik\workspace\IRstage1\src**
2. Compile the Java files:
 - `javac Porter.java TextParser.java Indexer.java QueryProcessor.java`.

2. Run the Text Parser

1. Run the TextParser.java file : open TextParser.java file as shown in the Picture below, then click on run option on top of the screen and then select start debugging.



Output:

- The output will be saved in a file named vsm_output.txt, performance_report.txt, parser_output.txt, forward_index.txt, inverted_index.txt in the project folder.
- Then Enter a term to search or 'quit' to exit: if the term you entered exists in any document it will show the document id and its frequency in that particular document.
- I also saved the document name and its mapped id to doc_ids.txt and word and its mapped id to word_ids.txt, these two text files are just for our reference so that we can easily understand the forward_index.txt and inverted_index.txt.