1. Create a ThreadPoolExecutor of size 20
2. Main thread reads file line by line and build an array list of certain size (i.e. 100 or 1000 lines)
3. Main thread creates a thread to process each array list of lines and put it in the thread pool executor for execution.
4. Each thread extracts words and word count from an array list of lines.
5. Each thread prepares an array list for MongoDB unordered bulk write operation which
6. Inserts “word” and “wordcount” in the MongoDB collection if does not exist OR
7. Updates “wordcount” by incrementing new count of words for a given word

e.g. UpdateOneModel<>(new Document(“word”,word),

new Document(“$inc”, new Document(“wordcount”,count)))

With Upsert true Update Options.

1. Each thread executes a bulk write operation to insert or update word count in the MongoDB collection within sharded cluster.