**ReadMe**

We designed and built an application that can reconstruct a PDF file to a shorter version that contains pages that are most relevant to the user's keyword input.

**Instruction**

**Design**

In this Swap Notes project, we’ve designed the following classes:

1.FileInputFilter, takes a file as input and filters the pages that meets condition.

2.AnalyzedPage, does vector space analysis or keyword frequency analysis for one page.

3.DocAnalyzer, filters documents based on user input, and user can select the sort order and the number of pages for output.

4.DocPrinter, saves the PDDocument as a pdf file to a directory that user selected.

5.GuiMain, represents the GUI (Graphical User Interface) for SwapNote

**CRC**

**Algorithm**

We implemented a ranking algorithm based on the vector space model to sort and filter each page by relevance to the input keyword list.

**API**