Project #5: NLP-based Document Summarization System	
Objective	Develop a system to automatically summarize long documents (e.g., legal contracts, research papers).
Functionality	• **Text Preprocessing:** Clean and tokenize document text. • **Summarization Model:** Use NLP models (e.g., BERT, GPT) to generate summaries. • **Summary Customization:** Allow users to specify summary length or focus areas. • **Export Options:** Export summaries in various formats (e.g., PDF, Word). • **Integration with Document Management Systems:** Automatically summarize documents stored in repositories.
Integrations	• Document management systems (e.g., SharePoint, Google Drive). • Cloud storage services (e.g., AWS S3, Google Cloud Storage). • Export tools for generating summaries in different formats.
Tasks	• Collect and preprocess document datasets. • Develop and fine-tune an NLP model for summarization. • Build a user interface for uploading documents and viewing summaries. • Integrate with document management systems (e.g., SharePoint, Google Drive). • Test and validate the system's accuracy and usability.
Development Constraints	• **Time:** 6 months. • **Budget:** $150,000.

SCOPE:

summarizing long documents using NLP and ML (legal documents, contracts, ect...)
the software needs to be able to process data it needs well, so it needs to learn on big data with low error %.
It needs to keep privacy of the data it has control to. Needs to be delivered in 6 months
Documents managing system

feasability:-

- tech requirements:
  - access to large amounts of data
  - NLP libraries
  - API
  - Cloud storage
  - Document management system
  - Strong domain

Human resources:

- 3 AI experts
- 4 Data analysist
- 1 Project manager
- 5 Developers

Chalanges:

- low error rate
- cpu-power for big data
- data leakage
- Big storage

Economic feasibility

- 300 K <Development cost>
- Data ⟶ 90k
- Testing & Development < 40 K >
- Total 430 K

Operation feasibility

- Organizational : ✓
- User: Wit the easy User interface ✓
  & wide range of operations
- System integration :
  The company's system has been using other similar AI tech

3 Work

- WBS :-

L1: NLP Processing Project

2 - PM
3 - RG
2 - Design & Prototyping
3 - NLP Model training
4 - Data collection
4 - D Pre Process
4 - D training

Cost

Size = 30 kLOC

PT = 3.67

PERS = high (0.8)

CPLX = medium (1.10)

Tool = high (0.90)

TEM = 0.8019

① effort = $3.67 \times (30)^{1.10} \times 0.8019 = 124$ Person-Months

② Cost = effort × Avrage Monthly sallary
= 124 × 23,000 = 2,852,000

Metric	Value
Project Size (KLOC)	3.67
Effort (Person-Months)	124
Cost Estimation (USD)	$2,852,000