Project on

“Text & String Tools using CustomTkinter”

UE25CS151A – PCPS

Submitted by:

Saharsh Varma PES2UG25CS465

Sahil Sinha PES2UG25CS468

Ryan Prince PES2UG25AM217

Rishi Arun PES2UG25CS724

Under the guidance of

Prof. Vanitha A N

PES University

Sep – Dec 2025

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

PES UNIVERSITY

(Established under Karnataka Act No. 16 of 2013)

Electronic City, Hosur Road, Bengaluru – 560 100, Karnataka, India

i. Problem Statement

Students and users frequently work with text stored in multiple file formats such as text files, Word documents, PDFs, and PowerPoint presentations. Extracting meaningful information like word frequency across these formats usually requires different tools or manual effort. Additionally, basic string-processing utilities such as palindrome checking and encryption/decryption are often taught theoretically without an interactive implementation.

The problem is to design a **single, user-friendly desktop application** that can: - Analyze textual content from multiple file formats - Perform common string operations interactively - Provide clear visual output with minimal user effort

---

ii. Approach / Methodology / Data Structures Used

Approach:

- A GUI-based desktop application was developed using CustomTkinter for a modern and responsive interface.
- The application is divided into three independent modules using a tab-based layout:
- Word Frequency Analyzer
- Palindrome Checker
- Caesar Cipher (Encryption/Decryption)

Methodology:

- Modular programming approach separating:
- Text-processing logic
- File-handling logic
- User interface logic
- Exception handling is used to manage unsupported files and empty inputs gracefully.

Data Structures Used

- Dictionary / Counter (collections.Counter): To store word-frequency mappings
- Lists: For collecting extracted text segments and characters
- Strings: For text manipulation and normalization

External libraries used: - `python-docx` for `.docx` files - `PyPDF2` for `.pdf` files - `python-pptx` for `.pptx` files

---

iii. Sample Input / Output

Word Frequency Tool

Input: - File: `sample.txt` - Content: `Hello world hello`

Output:

```
hello .................. 2
world .................. 1
```

Palindrome Checker

Input:

```
A man, a plan, a canal: Panama
```

Output:

```
Palindrome
```

Caesar Cipher

Input:

```
Text: ABC
Shift: 3
```

Output:

```
DEF
```

---

iv. Challenges Faced

- Extracting text reliably from different file formats
- Handling image-only or scanned PDFs where no text is available
- Maintaining aligned output formatting for word-frequency results
- Managing GUI responsiveness while processing large documents • Ensuring consistent behavior across different operating systems

---

v. Scope for Improvement

- Add OCR support for scanned PDFs and images
- Allow exporting results to text or CSV files
- Add options for sorting word frequency (alphabetical / top-N words)
- Include additional cryptographic algorithms
- Convert the application into a standalone executable
