**THE UNIVERSITY OF LAHORE**

*Department of Computer Science & IT Faculty of Information Technology*

Day Month Year

**DATE**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | **–** | 0 | 3 | **–** | 2 | 0 | 2 | 1 |

|  |  |
| --- | --- |
| *Write down the brief project topic, the title should not be confusing.* | |
| **PROJECT TITLE:** | Document Format Validator |
| **KEY WORDS:** |  |
| **DOMAIN OF THE PROJECT:** | Optical Character Recognition |
| **SUPERVISOR’S NAME:** | Sir Ijaz Khan Baig |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STUDENT INFORMATION** | | | |  |
| *Sr.* | *Student ID* | *Name* | *Email* | *Mobile* |
| 1. | 70062940 | ZOHAIB MEHMOOD | 70062940@student.uol.edu.pk | 0345-5520174 |
| 2. | 70062655 | ADIL BASHIR | 70062655[@student.uol.edu.pk](mailto:70066032@student.uol.edu.pk) | 0336-5046634 |

|  |
| --- |
| **PROBLEM STATEMENT** |
| In this day and age of digitization, our professional and non-professional lives are filled with documents whether they are assignments of a students, adds for recruitments, a proposal for business project or legally binding documents. All of these documents are being made electronically on computer or laptops. And the most important thing in all these types of documents is their format and styling (How data is presented). Although there are tools for such work but there are no tools that can check the formatting and styling of said documents in one simple click and highlight any abnormality. |
| **EXECUTIVE SUMMARY:**  OCR stands for "Optical Character Recognition." OCR is a technology that recognizes text within a digital image. It is commonly used to recognize text in scanned documents, but it serves many other purposes as well. OCR software processes a digital image by locating and recognizing characters, such as letters, numbers, and symbols. OCR is sometimes also referred to as text recognition.  We use this technique to cross check our document against desired format/style parameters.  This methodology can be used in various ways some of these are:   * It will provide an easy and simple method to proof check documents format. * It will highlight mistakes if any.   Can be used by anyone as it is available on the internet. |

|  |
| --- |
| **INTRODUCTION** |
| The goal of this project is to provide an easy way to check format/style of documents by making a platform where users can upload their .doc type files and check their formatting/styling against one of the pre-set parameters or parameters set by themselves.  At the end system will highlight any mistakes that don’t match the given parameters if they exist or display a message that the document is properly formatted against the given parameters. |

|  |
| --- |
| **COMPETITORS/COMPETITIVE ANALYSIS**  *.* |
| As of now our idea is unique, it has little to no competitors. As far according to my research there are some relevant work in on going but it’s not accurate enough to be utilized. |

|  |
| --- |
| **OBJECTIVES**  *.* |
| * System will read only .doc type files. * System will provide a range of parameters to be checked. * System will provide the ability to set customs parameters. * System will highlight anything that doesn’t match given parameters. * System will achieve minimum 95% accuracy. * Efficient and Accurate. |
| **MOTIVATION** |
| This very document and its sequels are the core motivation for this project. FYP is a must in every university all around the world and each and every university have their own specific style for format set for their FYP. This project provides an easy way to check the format and style standards to professors. This system can be used in real-time not only by professors but everyone who wants to make sure their document is properly formatted. |

|  |
| --- |
| **FEATURES OF PROJECT** |
| There are mainly only three features that matter, all of which are critical:   1. Allow users to upload their document file. 2. Provide a set of some famous pre-defined format parameters. 3. Allow users to make their own custom parameters to be compared against such as:    * Text Size: Size of text.    * Text Font: Font of text    * Text Color: Color of text e.g. black, white, red etc.    * Text Alignment: Text alignment whether if text should be left aligned, center or justified.    * Margin: Distance of text from either sides of the document.    * Line Spacing: Distance/spacing between two lines.    * Word Spacing: Distance/spacing between words.    * Header & Footer.   These are some of the parameters that can be custom set.   1. Highlight mistakes if any. |

|  |
| --- |
| **ARCHITECTURAL DESIGN** |
|  |

|  |
| --- |
| **IMPLEMENTATION TOOLS AND TECHNIQUES** |
| * WordPress * HTML * CSS * Office JS * OCR (Optical Character Recognition) |

|  |
| --- |
| **PROJECT PLAN:**  PART-I   1. Problem Identification. 2. Solution Proposal. 3. Implementing ML models. 4. SRS Document.   PART-II   1. Design Based Implementation. 2. Testing/Validation. 3. Project Report. 4. Deployment/Publishing. |
|  |

|  |
| --- |
| **REFERENCES** |
|  |

### Supervisor’s Signature: ………………………



**FYP Proposal Evaluation**

*(To be filled by students)*

|  |  |
| --- | --- |
| *Write down the brief project topic and should not be confusing.* | |
| **PROJECT TITLE:** | **Document Format Validator** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STUDENT INFORMATION**  *Write down the detail of all group members in BLOCK LETTERS ONLY.* | | | |  |
| *Sr.* | *Student ID* | *Name* | *Email* | *Mobile* |
| 1. | 70062940 | ZOHAIB MEHMOOD | 70062940@student.uol.edu.pk | 0345-5520174 |
| 2. | 70062655 | ADIL BASHIR | 70062655[@student.uol.edu.pk](mailto:70066032@student.uol.edu.pk) |  |

|  |  |
| --- | --- |
| **FOR EVALUATOR’S USE ONLY** | |
| Remarks: | Accepted with Minor changes |

**Suggested Improvements (***if any***):**

**Evaluated by:**

Name:

Signature:

Day Month

Year

**DATE**

**~~–~~**

**~~–~~**