**Spam/Junk Email Detector**

**Name:** Bryan Yadiel Caban Rodriguez

**Date Created:** Jan 13, 2025

**Program Description:**

A program that analyzes email messages to detect potential spam/junk content. It evaluates the message against a predefined list of spam keywords and phrases, calculates a spam score based on keyword occurrences, and determines the likelihood of the message being spam. The program provides detailed feedback including the spam score, likelihood rating, and any triggered keywords found in the message.

**Functions used in the Program:**

1. **Function Name:** calculate\_spam\_score(message, spam\_keywords)

**Description:** Analyzes an email message for spam keywords and calculates a spam score based on the number of matches found.

**Parameters:**

- message (str) - The email message to be analyzed

- spam\_keywords (list) - List of spam-related keywords and phrases

**Variables:**

- spam\_score (int) - Counter for the number of spam keyword matches

- triggered\_keywords (list) - List to store found spam keywords

- matches (list) - Stores regex matches for each keyword

**Logical Steps:**

1. Initialize spam score counter and triggered keywords list

2. For each spam keyword in the list:

- Create regex pattern with word boundaries for whole-word matching

- Find all case-insensitive matches in the message

- Add number of matches to spam score

- If matches found, add keyword to triggered list

3. Return the final score and list of triggered keywords

**Returns:** tuple (int, list) - The spam score and list of triggered keywords

2. **Function Name:** rate\_spam\_likelihood(spam\_score)

**Description:** Evaluates the likelihood of a message being spam based on its spam score.

**Parameters:**

- spam\_score (int) - The calculated spam score of the message

**Variables:** None

**Logical Steps:**

1. Compare spam\_score against predefined thresholds

2. Return appropriate likelihood rating:

- Score >= 10: "Very High"

- Score >= 5: "High"

- Score >= 2: "Medium"

- Score < 2: "Low"

**Returns:** str - The spam likelihood rating

3. **Function Name:** main()

**Description:** Main function that handles user interaction and coordinates the spam detection process.

**Parameters:** None

**Variables:**

- spam\_keywords (list) - Predefined list of spam-related keywords and phrases

- email\_message (str) - User-input email message to analyze

- spam\_score (int) - Calculated spam score from calculate\_spam\_score()

- triggered\_keywords (list) - List of found spam keywords

- spam\_likelihood (str) - Rating of spam likelihood

**Logical Steps:**

1. Define list of spam keywords

2. Prompt user for email message input

3. Call calculate\_spam\_score() to analyze the message

4. Call rate\_spam\_likelihood() to determine spam probability

5. Display results:

- Spam score

- Spam likelihood

- List of triggered keywords (if any)

**Returns:** None

**Overall Program Logical Flow:**

1. Program starts by running main()

2. User inputs email message for analysis

3. Message is processed through calculate\_spam\_score():

- Checks for keyword matches

- Counts occurrences

- Tracks triggered keywords

4. Spam score is evaluated by rate\_spam\_likelihood()

5. Results are displayed to user:

- Final spam score

- Likelihood rating

- Any triggered keywords found

**Link to your repository:** https://github.com/xXTeinsXx/COP2373