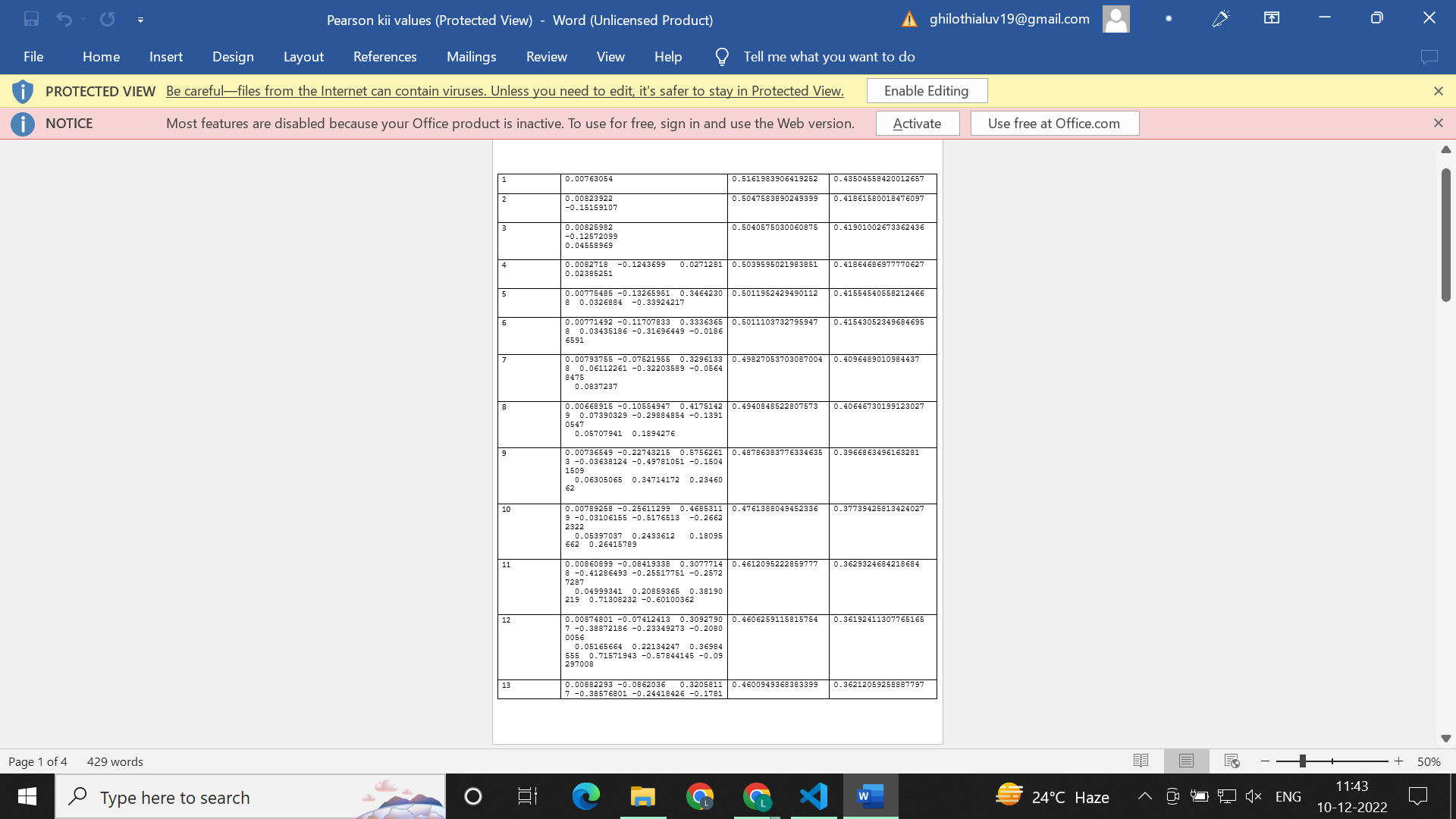
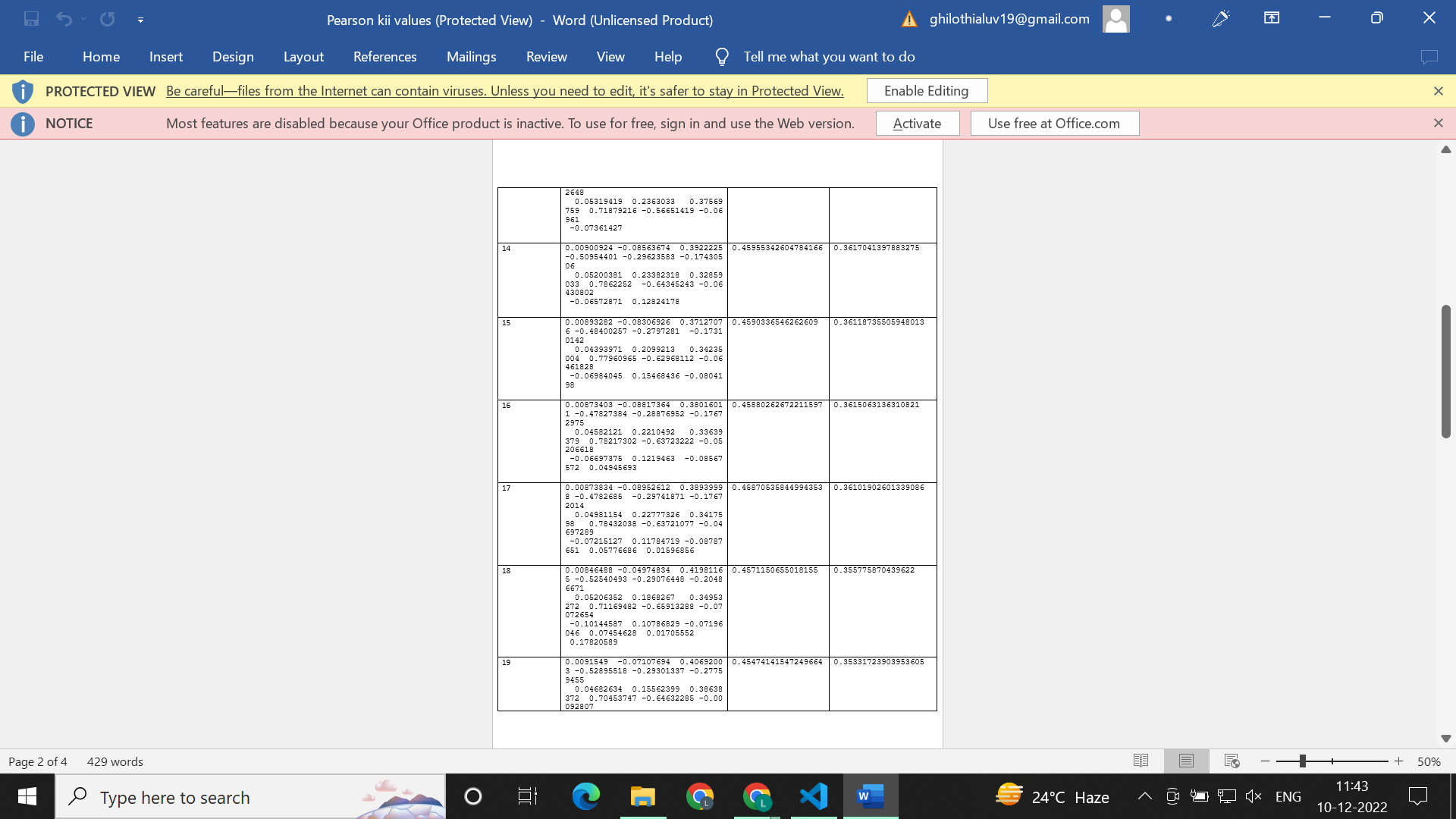
**PEARSON’S CORRELATION**

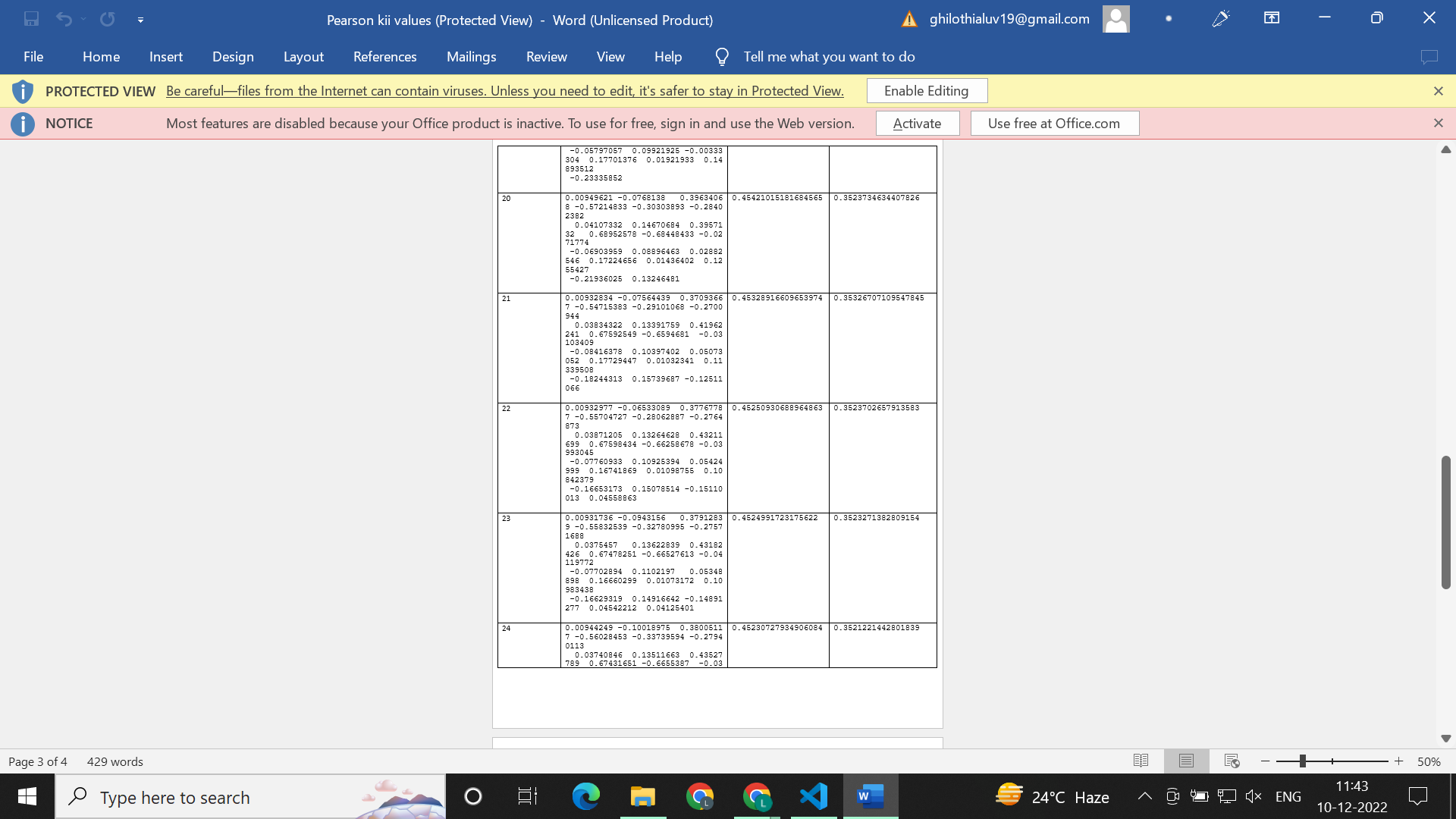
**The Pearson correlation coefficient (*r*) is the most common way of measuring a linear correlation. It is a number between –1 and 1 that measures the strength and direction of the relationship between two variables.**

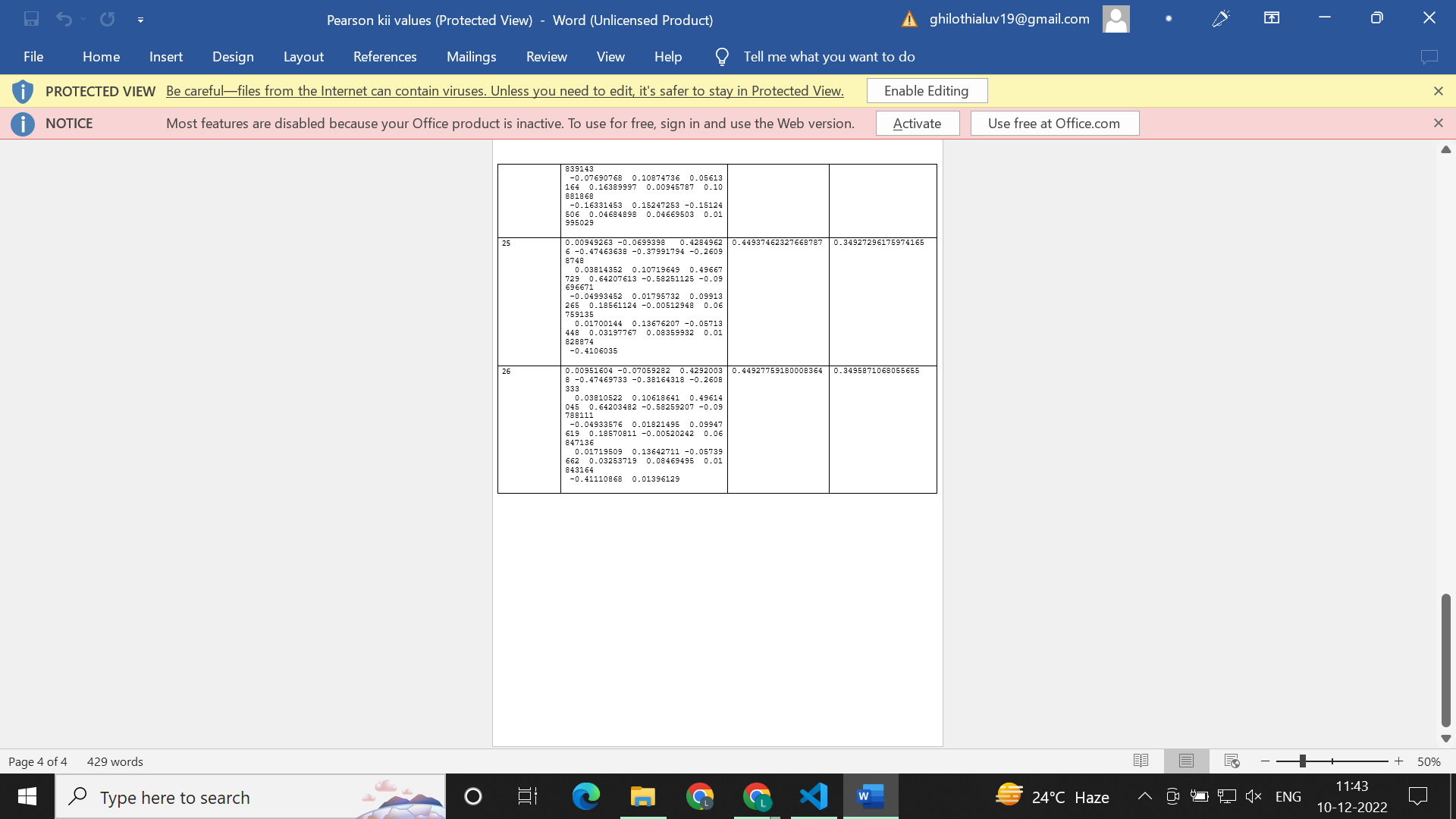
**The Pearson correlation coefficient is a good choice when all of the following are true:**

* **The relationship is linear: “Linear'' means that the relationship between the two variables can be described reasonably well by a straight line. You can use a scatter plot to check whether the relationship between two variables is linear.**

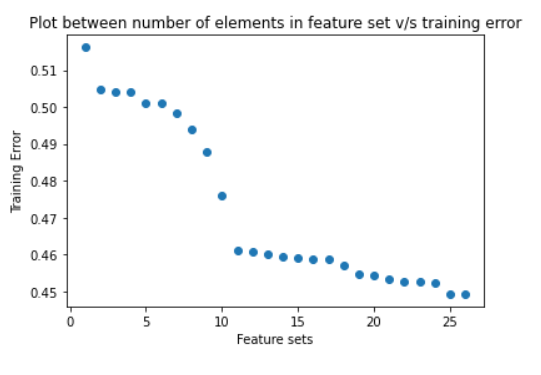




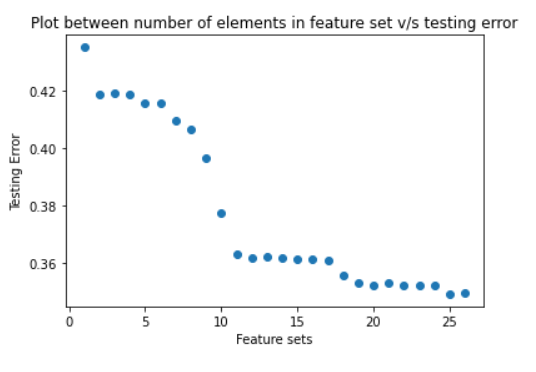




**With the increasing number of models in this method, both the training and testing errors decrease.**

****

**In the plot between number of elements in feature set vs training error, the training error decreases with the increasing feature sets.**

****

**In the plot between number of elements in feature set vs testing error, the testing error decreases with the increasing feature sets.**

**As it is evident from the graph above, if we don’t want to select the complete feature set, we see a minima in a feature set of 20 most important features.**