Sure! Here are detailed explanations for each of your questions:

### 1. P-value and its Significance

\*\*P-value\*\* is a statistical measure that helps researchers determine whether their hypotheses are correct. In hypothesis testing, the p-value is used to determine the significance of the results.

- \*\*Significance\*\*: The p-value measures the probability that the observed data (or something more extreme) would occur if the null hypothesis were true. A smaller p-value indicates stronger evidence against the null hypothesis.

- \*\*Threshold\*\*: Commonly, a p-value threshold (α level) of 0.05 is used. If the p-value is less than 0.05, the results are considered statistically significant, meaning there's strong evidence against the null hypothesis.

- \*\*Calculation\*\*: The p-value is calculated using statistical tests like t-tests, chi-square tests, etc., based on the distribution of the test statistic under the null hypothesis.

### 2. Z-test vs. t-test

Both Z-tests and t-tests are used to determine if there is a significant difference between sample means.

- \*\*Z-test\*\*:

- \*\*When to use\*\*: When the sample size is large (typically n > 30) and the population variance is known.

- \*\*Difference\*\*: Uses the Z-distribution (normal distribution).

- \*\*t-test\*\*:

- \*\*When to use\*\*: When the sample size is small (n ≤ 30) or the population variance is unknown.

- \*\*Difference\*\*: Uses the t-distribution, which accounts for additional uncertainty in small samples.

### 3. TFIDF vs. Word2Vec

\*\*TFIDF (Term Frequency-Inverse Document Frequency)\*\* and \*\*Word2Vec\*\* are both techniques for representing text data but work very differently.

- \*\*TFIDF\*\*:

- \*\*How it works\*\*: Measures the importance of a word in a document relative to a collection of documents (corpus).

- Term Frequency (TF): Frequency of a word in a document.

- Inverse Document Frequency (IDF): Measures how common or rare a word is across the corpus.

- \*\*Representation\*\*: Produces sparse vectors representing the importance of words in documents.

- \*\*Use case\*\*: Good for traditional text classification tasks.

- \*\*Word2Vec\*\*:

- \*\*How it works\*\*: Uses neural networks to generate dense vector representations of words based on their context in sentences.

- \*\*CBOW (Continuous Bag of Words)\*\*: Predicts a target word from its context.

- \*\*Skip-gram\*\*: Predicts context words from a target word.

- \*\*Representation\*\*: Produces dense vectors capturing semantic relationships between words.

- \*\*Use case\*\*: Suitable for capturing semantic meaning and analogies.

### 4. Adjusted R²

\*\*Adjusted R²\*\* is a modified version of R² that adjusts for the number of predictors in the model.

- \*\*Purpose\*\*: Adjusted R² accounts for the number of predictors, preventing overfitting. Unlike R², it can decrease if additional predictors do not improve the model.

- \*\*Calculation\*\*: Adjusted R² = 1 - [(1 - R²)(n - 1) / (n - p - 1)], where n is the number of observations, and p is the number of predictors.

### 5. Model Predictions Beyond the Target Range

For predicting outcomes between 0 and 5, \*\*linear models\*\* like \*\*Linear Regression\*\* and \*\*Polynomial Regression\*\* might predict values outside this range due to their continuous nature.

- \*\*Linear Regression and Polynomial Regression\*\*: Can generate predictions beyond 5 because they fit a continuous function to the data.

- \*\*Random Forest and XGBoost\*\*: These models can be more constrained by the range of training data but could also extrapolate beyond the training range if not explicitly bounded.

### 6. Word2Vec vs. GloVe

\*\*Word2Vec\*\* and \*\*GloVe\*\* are both word embedding techniques but differ in their approaches.

- \*\*Word2Vec\*\*:

- \*\*How it works\*\*: Uses a neural network to predict target words from context (CBOW) or context from target words (Skip-gram).

- \*\*Training\*\*: Trained on local context of words within a window.

- \*\*GloVe (Global Vectors for Word Representation)\*\*:

- \*\*How it works\*\*: Uses matrix factorization techniques on the word co-occurrence matrix to find word vectors.

- \*\*Training\*\*: Considers global word-word co-occurrence statistics from a corpus.

- \*\*Key Difference\*\*: Word2Vec focuses on local context, while GloVe leverages global context from the entire corpus to create word embeddings.