FROM CHATGPT:

STANDARDIZATION USED IN THE BELOW ALGORITHMS:

Standardization is often necessary for machine learning algorithms that rely on distance-based metrics or optimization methods. The reason for standardization is to ensure that all features contribute equally to the model fitting process and prevent one feature from dominating due to its scale. Here are some machine learning algorithms that commonly benefit from standardization:

1. **K-Nearest Neighbors (KNN):**
   - KNN calculates distances between data points to make predictions.
   - Standardizing features ensures that the scale of each feature does not disproportionately impact the distance calculations.
2. **Support Vector Machines (SVM):**
   - SVM aims to find the hyperplane that best separates classes.
   - Standardization is crucial to make sure that all features have similar scales, preventing features with larger magnitudes from dominating the decision boundary.
3. **Principal Component Analysis (PCA):**
   - PCA identifies the principal components by maximizing variance.
   - Standardization is important to give equal weight to all features during the variance-capturing process.
4. **Gradient Descent-Based Algorithms:**
   - Algorithms like linear regression, logistic regression, and neural networks that use gradient descent are sensitive to feature scales.
   - Standardization ensures faster convergence during optimization.
5. **K-Means Clustering:**
   - K-Means relies on distance metrics to assign data points to clusters.
   - Standardization prevents features with larger scales from having a disproportionate impact on the clustering process.
6. **Linear Discriminant Analysis (LDA):**
   - LDA, similar to PCA, involves calculating covariance matrices and eigenvectors.
   - Standardization ensures that the covariance matrices are not dominated by features with larger scales.
7. **Neural Networks:**
   - Neural networks often use optimization algorithms like gradient descent.
   - Standardization ensures that the learning process is efficient and features contribute equally to the weight updates.

8. **Ridge and Lasso Regression:**
    - Regularized regression techniques like Ridge and Lasso are sensitive to feature scales.
    - Standardization helps in preventing regularization from being biased toward features with larger magnitudes.

Remember, not all algorithms require standardization. Tree-based algorithms like Decision Trees and Random Forests, for instance, are generally insensitive to feature scales. Always check the documentation or experiment with your specific dataset to determine the impact of standardization on the performance of your chosen algorithm.

STANDARDIZATION WON'T USED IN THE BELOW ALGORITHMS:

Standardization is not always necessary or beneficial for all machine learning algorithms, particularly those that are not sensitive to the scale of input features. Here are some examples of algorithms where standardization is typically not used or has a minimal impact:

1. **Decision Trees:**
    - Decision Trees and ensemble methods like Random Forests and Gradient Boosted Trees are generally insensitive to feature scales.
    - These algorithms make decisions based on feature thresholds, and the scale of features does not affect the splits.
2. **Random Forests:**
    - Random Forests, being an ensemble of Decision Trees, are not sensitive to feature scales.
    - Each tree in the forest independently makes decisions based on the features without considering their scales.
3. **Gradient Boosted Trees:**
    - Gradient Boosted Trees, like Random Forests, are robust to feature scales.
    - The boosting process focuses on reducing errors, and the individual trees do not rely on feature scales.
4. **Naive Bayes:**
    - Naive Bayes classifiers assume independence between features given the class, and they are not affected by feature scales.
5. **k-Means Clustering (with Euclidean Distance):**
    - While k-Means is sensitive to distance metrics, if Euclidean distance is used and the clusters are spherical, then standardization might not be crucial. However, it depends on the specific characteristics of the data.
6. **Nearest Neighbors (with custom distance metrics):**

- If custom distance metrics are used in k-Nearest Neighbors (KNN), standardization might have different effects. However, with the default Euclidean distance, standardization is typically recommended.

7. **Association Rule Learning (Apriori, Eclat):**
   - Algorithms like Apriori and Eclat for association rule learning do not rely on the scale of individual features.
8. **Ordinal Regression:**
   - Ordinal regression models, which deal with ordinal target variables, often do not require standardization.

It's important to note that while these algorithms may not require standardization for their core functionality, there could be exceptions based on specific characteristics of the dataset or the problem at hand. Additionally, when in doubt, it's a good practice to experiment with both standardized and non-standardized versions of the data to observe the impact on algorithm performance.