**CMPE-257 Week - 5 Notes**

Submitted By :

Siddhant Sancheti(016710421)

**Naive Bayes**

* Naive Bayes is a popular machine learning algorithm for classification tasks
* It is based on Bayes' theorem which states that the probability of an event (A) given the occurrence of another event (B) can be calculated as the product of the probability of event B given A, and the probability of event A, divided by the probability of event B
* The "naive" in Naive Bayes refers to the assumption that the features used for classification are independent of each other, which is often not true in practice
* Despite this assumption, Naive Bayes can still perform well in practice, especially in cases where the dimensionality of the feature space is high
* Naive Bayes can handle both binary and multi-class classification tasks
* There are three main types of Naive Bayes classifiers: Gaussian Naive Bayes, Multinomial Naive Bayes, and Bernoulli Naive Bayes, which are suited for different types of data
* Gaussian Naive Bayes assumes that the features follow a Gaussian distribution, and is often used for continuous data
* Multinomial Naive Bayes is used for discrete data with multiple classes, such as text classification
* Bernoulli Naive Bayes is used for binary data, such as spam detection
* Naive Bayes can be trained quickly and requires relatively little data compared to other classification algorithms
* However, its performance may suffer when the assumption of feature independence is strongly violated, or when there is little training data available
* Naive Bayes is widely used in applications such as spam filtering, sentiment analysis, and recommendation systems.

**Logistic Regression versus Decision Trees**

* Logistic regression and decision trees are two popular algorithms for classification tasks in machine learning
* Logistic regression is a parametric algorithm that models the relationship between the independent variables (features) and the dependent variable (target) using a logistic function
* Decision trees, on the other hand, are non-parametric algorithms that create a tree-like structure to recursively partition the feature space based on the values of the features
* Logistic regression is suited for tasks where the relationship between the features and the target is linear, and where the decision boundary is a hyperplane
* Decision trees are suited for tasks where the relationship between the features and the target is non-linear, and where the decision boundary is non-linear and can be approximated by a piecewise constant function
* Logistic regression is less prone to overfitting than decision trees, especially when the number of features is large compared to the number of training examples
* Decision trees can be prone to overfitting, especially when the depth of the tree is large or when the training data is noisy
* Logistic regression can be interpreted easily, as the coefficients of the logistic function can be used to determine the importance and directionality of each feature in the classification decision
* Decision trees can also be interpreted easily, as the branches and nodes of the tree can be used to determine the importance and interaction of each feature in the classification decision
* Logistic regression is computationally efficient and can be trained quickly even on large datasets
* Decision trees can be computationally expensive to train, especially when the depth of the tree is large or when the number of features is large
* Logistic regression is widely used in applications such as credit scoring, churn prediction, and fraud detection
* Decision trees are widely used in applications such as customer segmentation, product recommendation, and medical diagnosis.

**Ensemble Learning Methods**

* Ensemble learning is a machine learning technique that combines multiple models to improve prediction accuracy and reduce overfitting
* Ensemble learning methods can be divided into two categories: bagging and boosting
* Bagging methods involve training multiple models on different subsets of the training data and then combining their predictions to obtain a final prediction
* Random Forest is a popular bagging algorithm that trains multiple decision trees on random subsets of the training data and then combines their predictions using a majority vote
* Bagging methods can improve prediction accuracy and reduce overfitting by reducing the variance of the models
* Boosting methods involve training multiple models sequentially, with each model focusing on the examples that were misclassified by the previous model
* Gradient Boosting is a popular boosting algorithm that trains decision trees sequentially, with each tree trying to correct the errors made by the previous tree
* Boosting methods can improve prediction accuracy by reducing both bias and variance of the models
* Ensemble learning methods can be combined with other machine learning techniques, such as feature selection and hyperparameter tuning, to further improve performance
* Stacking is another ensemble learning method that involves training multiple models and then using a meta-model to combine their predictions
* Stacking can improve prediction accuracy and reduce overfitting by learning to weigh the predictions of the individual models
* Ensemble learning methods can be computationally expensive, especially when the number of models or the size of the training data is large
* Ensemble learning methods are widely used in applications such as image classification, natural language processing, and fraud detection.

**An Introduction to Restricted Boltzmann Machines**

* Restricted Boltzmann Machines (RBMs) are generative models that can learn to represent complex distributions of data
* RBMs are composed of visible units and hidden units, and the connections between them are undirected
* RBMs can be trained using the Contrastive Divergence algorithm, which involves updating the weights of the connections to maximize the likelihood of the training data
* RBMs can be used for a variety of tasks, including feature learning, dimensionality reduction, and recommendation systems
* RBMs can be extended to Deep Belief Networks (DBNs), which are composed of multiple layers of RBMs
* DBNs can be trained using the greedy layer-wise pretraining algorithm, which involves training each layer of RBMs in a unsupervised manner and then fine-tuning the entire network using backpropagation
* DBNs can be used for a variety of tasks, including image recognition, speech recognition, and natural language processing
* RBMs and DBNs can be used for unsupervised learning, semi-supervised learning, and supervised learning
* RBMs and DBNs can be computationally expensive, especially when the number of visible units or hidden units is large
* RBMs and DBNs are widely used in research and industry, including applications such as recommendation systems, drug discovery, and image recognition

**Articles**

* Naive Bayes - <https://towardsdatascience.com/how-to-split-a-dataframe-into-train-and-test-set-with-python-eaa1630ca7b3>
* Logistic Regression versus Decision Trees -

<https://blog.bigml.com/2016/09/28/logistic-regression-versus-decision-trees/>

* Ensemble Learning Methods -

<https://towardsdatascience.com/simple-guide-for-ensemble-learning-methods-d87cc68705a2>

* An Introduction to Restricted Boltzmann Machines -

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.390.9412&rep=rep1&type=pdf>