Comparison of Algorithms

|  |  |  |  |
| --- | --- | --- | --- |
| ALGORITHM | ADVANTAGE | DISADVANTAGE | APPLICATION |
| Linear Regression | Simple algorithm,  helps in finding relationship intuitively | It oversimplifies the problem and always assume a linear relationship between variables | Stock market prediction ,  customer survey results analysis,  weather forecasting |
| SVM | Due to regularization of parameters over fitting problem is avoided, kernel trick helps svm to solve complex problem. SVM is not solved for local optima. | Choosing a good kernel is not easy, Difficult to understand and decode the final model, hence tough for business logic. | Prediction problems, Handwriting Recognition, |
| Decision tree | Very useful in classification and decision problems , it is capable of handling missing values in attributes, | Over fitting of data , sampling error, it achieve local optimum may not global, difficult to control size | Customer Relationship Management ,  Fraudulent Statement Detection  health area |
| Naive  Bayes Classification | it will converge quicker than discriminative models like logistic regression, Need less training data, Not sensitive to irrelevant features | makes a very strong assumption on the shape of your data distribution, continuous features creates a problem and a large amount of data is lost | Real-time Prediction,  Text classification Spam Filtering Sentiment Analysis  Recommendation system |
| LOGISTIC REGRESSION | It gives the information of how relevant the prediction is and direction of prediction | Independent observation is required , over fitting, limited outcome variable | If you want to predict any classification problem |
| K-means clustering | Easy to implement  k-Means may produce tighter clusters  An instance can change cluster i.e. can move to another cluster when the centroid are computed | Initial guesses have a strong impact on the final results  Sensitive to scaling i.e. on normalizing the data set result completely changes. | document clustering  image compression  News on sites  Color prediction |
| GAUSSIAN MIXTURE MODEL | GMM is a lot more flexible  Produce valid parameters for the mixture distribution | Converges to local optimum  Less relevant in 1- D | Anomaly detection |