BYTE QUEST

Vasavi College of Engineering

Department of Computer Science and Engineering

August 30, 2019

Volume 71

Contents:

- ∗ KNN ALGORITHM

- ∗ SVM ALGORITHM

- ∗ LOGISTIC REGRESSION

Byte Quest is the article published by the CSE dept of Vasavi College of Engineering regarding the latest innovative Technologies and Software that have been emerged in the competitive world. The motto of this article is to update the people regarding the improvement in technology. The article is designed by the active participation of students under the guidance of faculty coordinators.

 Good, bad or indifferent if you are not investing in new technology, you are going to be left behind.

-Philip Green

 Once a new technology rolls over you, if you're not part of the steamroller, you're part of the road.

-Stewart Brand

FACULTY CO-ORDINATORS

M.SUNDARI (ASST. PROFESSOR)

GARIMA JAIN(ASST. PROFESSOR)

STUDENT COORDINATORS

|	
NIKITHA (4/4 CSE-A)	ABHINAV (4/4 CSE-B)
G. CAROL EUNICE (3/4 CSE-A)	APARNA DARBHA(3/4 CSE-B)
ABHINAV (2/4 CSE-A)	ANISHA (2/4 CSE-B)

K-NEAREST NEIGHBOURS

KNN's main disadvantage of becoming significantly slower as the volume of data increases makes it an impractical choice in environments where predictions need to be made rapidly. Moreover, there are faster algorithms that can produce more accurate classification and regression results. However , provided you have sufficient computing resources to speedily handle the data you are using to make predictions, KNN can still be useful in solving problems that have solutions that depend on identifying similar objects. An example of this is using the KNN algorithm in recommender systems, an application of KNN-search.

At scale, this would look like recommending products on Amazon, articles on Medium, movies on Netflix, or videos on YouTube.
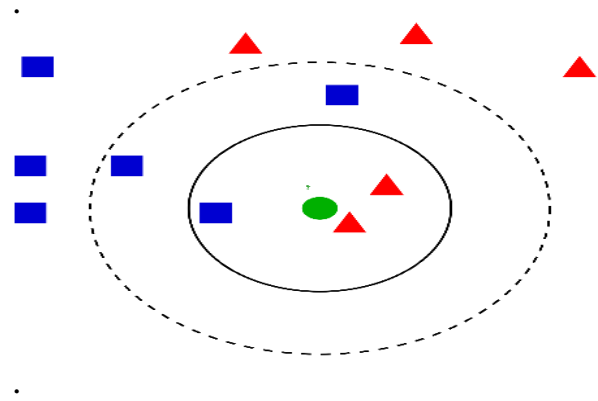
VISHWANATH (CSE –B 2/4)

SUPPORT VECTOR MACHINE
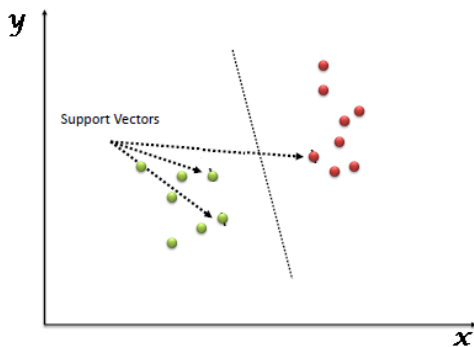
Support Vector Machine" (SVM) is a supervised algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes (hyper-plane/ line).

In SVM, it is easy to have a linear hyper-plane between these two classes. But, another burning question which arises is, should we need to add this feature manually to have a hyper-plane. No, SVM has a technique called the **kernel trick**. These are functions which takes low dimensional input space and transform it to a higher dimensional space. It is mostly useful in non-linear separation problem. Simply put, it does some extremely complex data transformations, then find out the process to separate the data based on the labels or outputs you've defined. It converts not separable problem to separable problem, these functions are called kernels.

VAISHNAVI (CSE B 2/4)

LOGISTIC REGRESSION

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, **logistic regression** is estimating the parameters of a logistic model (a form of binary regression). Mathematically, a binary logistic model has a dependent variable with two possible values, such as pass/fail which is represented by an indicator variable, where the two values are labeled "0" and "1". In the logistic model, the log-odds (the logarithm of the odds) for the value labeled "1" is a linear combination of one or more independent variables ("predictors"); the independent variables can each be a binary variable (two classes, coded by an indicator variable) or a continuous variable (any real value). The corresponding probability of the value labeled "1" can vary between 0 (certainly the value "0") and 1 (certainly the value "1"), hence the labeling; the function that converts log-odds to probability is the logistic function, hence the name. The unit of measurement for the log-odds scale is called a *logit*, from *logistic unit*, hence the alternative names.

The binary logistic regression model has extensions to more than two levels of the dependent variable: categorical outputs with more than two values are modeled by multinomial logistic regression, and if the multiple categories are ordered, by ordinal logistic regression, for example the proportional odds ordinal logistic model.[1] The model itself simply models probability of output in terms of input, and does not perform statistical classification (it is not a classifier), though it can be used to make a classifier, for instance by choosing a cutoff value and classifying inputs with probability greater than the cutoff as one class, below the cutoff as the other; this is a common way to make a binary classifier.

$$\text{sig}(t) = \frac{1}{1+e^{-t}}$$

NIDHI (CSE-A 3/4)