1 Classification

1. Derive the dual formulation of the soft support vector machine algorithm.
2. Derive the Discriminant function of the Quadratic Discriminant Analysis.
3. Derive the Gaussian process classifier by using the Laplace approximation.
4. For the following data, build a tree by using: a) Gini index, b) the depth of tree is two. The target variable is **Loan Approved**.

Table 1: Loan Approval Dataset for Tree Algorithms

Age	Income	Credit Score	Employment Type	Education	Home Ownership	Previous Loan	Married	Loan Approved
32	45,213	689	Full-time	Bachelor	Rent	No	Yes	Yes
45	78,234	734	Self-employed	Master	Own	Yes	Yes	Yes
28	31,245	612	Part-time	High School	Rent	No	No	No
51	92,345	756	Full-time	PhD	Mortgage	No	Yes	Yes
36	42,367	589	Full-time	Bachelor	Rent	Yes	Yes	No
42	67,342	698	Self-employed	Master	Own	No	No	Yes
29	28,765	543	Part-time	High School	Rent	Yes	No	No
58	84,567	812	Full-time	Bachelor	Own	No	Yes	Yes
31	39,234	654	Full-time	Bachelor	Mortgage	No	Yes	Yes
47	72,389	723	Self-employed	Master	Own	Yes	Yes	Yes
26	25,678	598	Unemployed	High School	Rent	No	No	No
39	56,321	687	Full-time	Bachelor	Mortgage	Yes	Yes	Yes
53	89,234	789	Full-time	Master	Own	No	Yes	Yes
34	42,345	623	Part-time	Bachelor	Rent	No	No	No
44	67,345	712	Full-time	Master	Mortgage	No	Yes	Yes
27	29,876	567	Unemployed	High School	Rent	Yes	No	No
49	81,234	745	Self-employed	PhD	Own	No	Yes	Yes
37	48,765	678	Full-time	Bachelor	Rent	Yes	Yes	Yes
41	52,345	701	Full-time	Master	Mortgage	No	No	Yes
30	34,567	634	Part-time	High School	Rent	No	Yes	No

5. For the data set in https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic, try following algorithms learned in the class: a) Discriminate function (mse, ppn, svm); b) Generative method (LDA, QDA, Naive Bayes (Gaussian)); c) logistic regression, neural network; d) KNN, tree method, random forest, e): kernel method (kernel svm, Gaussian process classifier). Report the parameters of your algorithm, and the accuracy on the test set.