INFORMATION THEORY AND STATISTICAL LEARNING GROUP

Lab Convener: Dr. Ambedkar Dukkipati

Computer Science and Automation

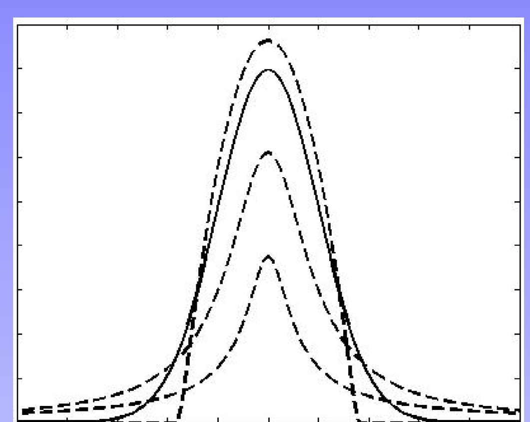
Indian institute of Science

Properties of Tsallis Divergence Minimization

Exponential distribution obtained from KL divergence minimization

generalization

Power law distribution obtained from Tsallis divergence minimization



—— Exponential distribution

----- Power law distributions

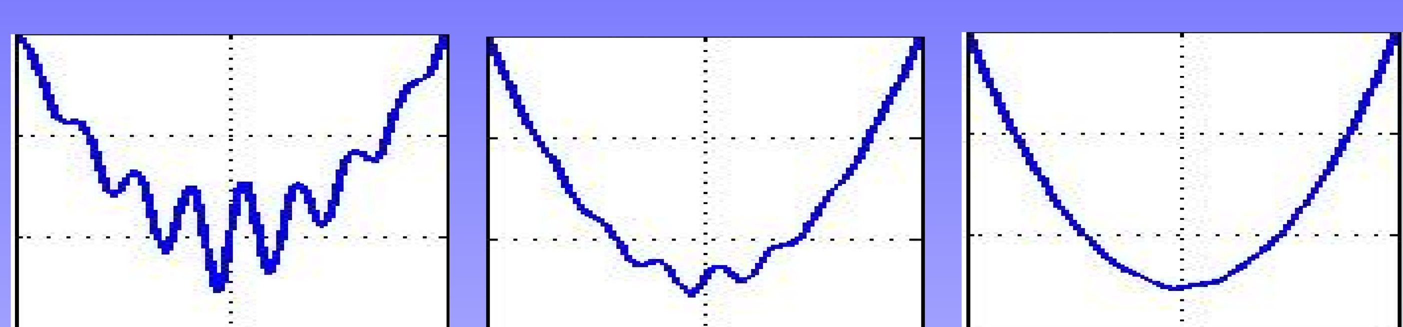
Our Work:

Generalize the important properties of Kullback-Leibler divergence minimization, such as Subset Independence, Pythagorean property for Tsallis divergence minimization.

Applications of Non-extensive Information Theory

q-Gaussian distribution → Generalization of normal distribution

q-Gaussians applied to Function Smoothening



Unsmoothed | Gaussian Smoothed | q-Gaussian Smoothed

q-Gaussians applied to Clustering and Classification

Positive definiteness of q-Gaussian kernel

q-SVM formalization

q-Gaussian kernel K-means

q-Gaussian Processes

Feature Selection from gene expression

Aim: To select a subset of genes that can maximize the discrimination between various phenotypes.

Given

A set of gene expression profiles and the corresponding phenotypes.

Method

Each gene expression profile forms a data item with the corresponding phenotype as the label.

The features are the genes .The corresponding gene expression denotes the value of feature.

Apply feature selection using maximum discrimination to obtain the set of features that maximize discrimination.

Label Ranking using Minimum Description Length Principle

Aim: To rank the labels according to their relevance to a given data item.

Intuition

More the regularity in data, more it can be compressed.

Learning ↔ Finding regularity in data.

Algorithm

Feature Selection using Maximum Discrimination

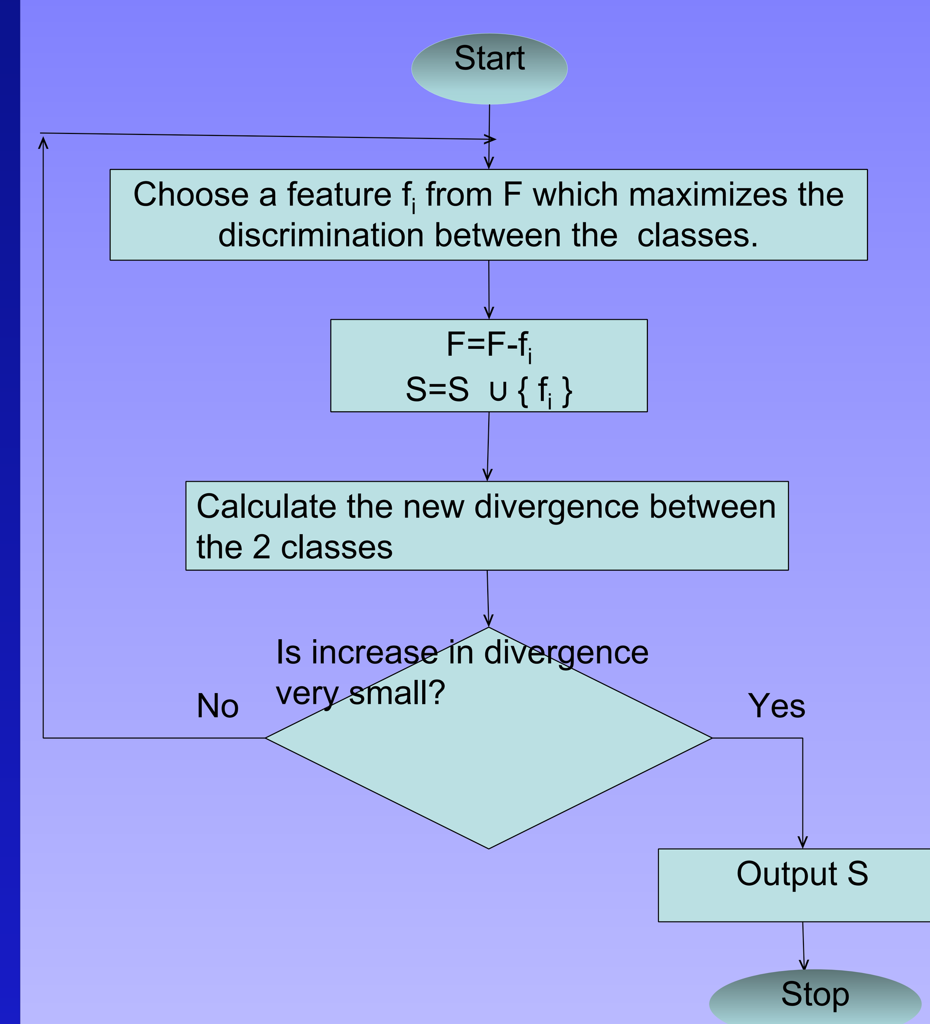
Aim: To select the minimal set of features which maximize the discrimination between classes.

Given

$F=<f_1,f_2,\ldots,f_n>$ → Given set of features

$S=\{\ \}$ → Selected features

Flowchart

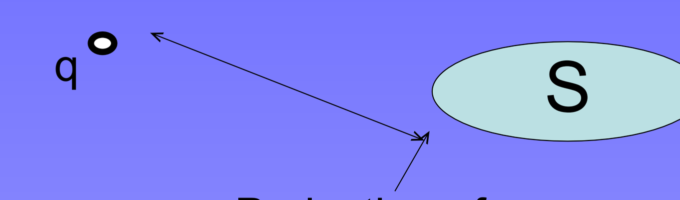


Ref: *Maximum entropy and maximum discrimination classifiction with Jefferys divergence*

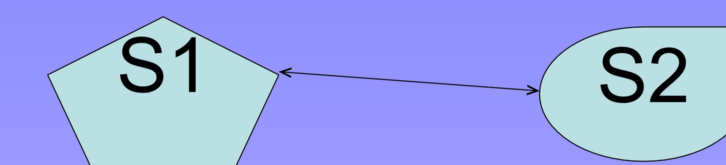
Ambedkar Dukkipati, Abhay Kumar Yadav, M. Narasimha Murty

Jensen Shannon Divergence Minimization

Jensen Shannon Projection and its properties

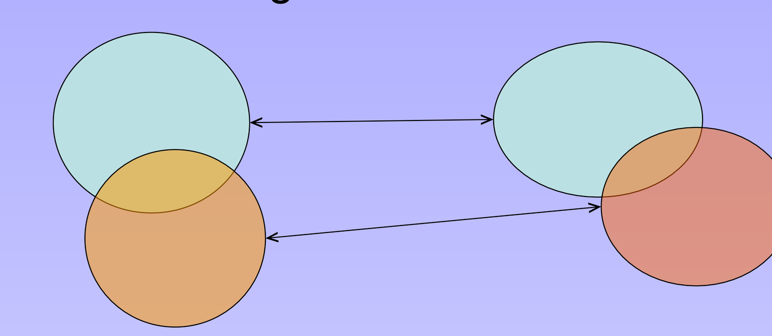


Jensen Shannon Divergence between 2 sets



Application to maximum disrimination classification

- In the algorithm for feature selection using maximum discrimination, Jensen Shannon divergence can be used to calculate divergence between the classes



LAB MEMBERS

***Jithin Vachery**, PhD., Computer Science & Automation*

***Anirban Sinha**, M.E., Computer Science & Automation*

***Gaurav Pandey**, M.E., Computer Science & Automation*

***D Manju Venkat Sai Sriram**, M.E., Computer Science & Automation*

***Debarghya Ghoshdastidar**, M.E., System Science & Automation*