PRML Course Project

Team Member – Hritik Kumar, Arpit Agarwal, Akshat Karani

For the PRML course project, we have chosen the Topic Modeling and Latent Dirichlet Allocation (LDA) as our problem statement.

What is LDA?

Topic Modeling is a type of statistical modeling for discovering the abstract "topics" that occur in a collection of documents. **Latent Dirichlet Allocation** (LDA) is an example of topic model and is used to classify text in a document to a particular topic.

In this project, we are going to apply LDA to a set of documents and split them into topics.

Implementation Plan

We will train our LDA model using "iitdh" broadcast mail and news article datasets available on kaggle.

A full implementation plan is yet to be finalised.