

INTRODUCTION AND APPLICATIONS OF MACHINE LEARNING



Contacts

Instructor

- Shreyas Bhat - 9082080984
- S I Harini - 7021247073
- Rishav Mukherji - 9920741703

Mentor

- Hardik Shah - 9427925103

Colab Notebooks

- Pandas
- Matplotlib

Pandas

Why Pandas?

NumPy is great. But it lacks a few things that are conducive to doing statistical analysis. By building on top of NumPy, pandas provides

- Labeled arrays
- Heterogeneous data types within a table
- "Better" at handling missing data
- Convenient methods (groupby, rolling, resample)
- More data types (Categorical, Datetime)

Data visualization using matplotlib

Sometimes data does not make sense until you can look at in a visual form, such as with charts and plots.

Being able to quickly visualize your data samples for yourself and others is an important skill both in applied statistics and in applied machine learning.

Data visualization is useful for data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modeling output, and presenting results.

In this lecture, we will learn to use matplotlib to visualize data effectively.

THANK YOU!

