HELP YELP
Machine Learning on Yelp Dataset

Team Members – Bhavna Menghrajani, Kunal Singh Deora, Manasi Laddha,
Sarthak Agarwal, Swarna Satishkumar Dommeti

Project Idea–

We are using Yelp dataset for our big data project to build a predictive model that will use the data having various businesses entities like users, reviews (rating), tips, check-ins, photos, etc. in an innovative way to help yelp in better analyzing their businesses in terms of cultural trends, seasonal trends, identifying trend setters and inferring businesses categories.

Currently, yelpers provide the review for a business on various grounds. It is not only the text but also the image that contributes to a review. The type of image content on Yelp is determined by its Image Classifier and classified as Food, Menu, Inside, Outside and Drinks. However, Yelp is unable to leverage this data to help the businesses expand to targeted audience. For example, what cuisines are Yelpers raving about in these different countries? Or Are there more reviews for sports bars on major game days? If so, Yelp can leverage the answers for these questions using ours machine learning based predictive model.

The Dataset Features –
* **2.7M** reviews and **649K** tips by **687K** users for **86K** businesses
* **566K** business attributes, e.g., hours, parking availability, ambience.
* Social network of **687K** users for a total of **4.2M** social edges.
* Aggregated check-ins over time for each of the **86K** businesses
* **200,000** pictures from the included businesses

Team Members –

Member Name	Work Responsibility/ Distribution
Bhavna, Sarthak, Manasi	Research applicable Machine Learning Algorithms, Frameworks, Libraries and Technical Papers
Kunal, Swarna	Featuring Engineering – Data cleaning, data conversation and feature analysis
Bhavna, Sarthak, Manasi, Kunal, Swarna	Designing our Predictive Model Architecture of data analysis
Bhavna, Kunal, Sarthak	Setting up the cloud cluster
Manasi, Swarna, Sarthak	Developing and Testing the neural networks and machine learning techniques used in the model

Bhavna, Kunal, Manasi, Sarthak	Deploy the solution on code and tune the accuracy
Swarna, Manasi, Bhavna, Sarthak, Kunal	Virtualization of the data, documentation and presentation