# Hackathon Instructions: Cross-Sell Prediction

## Objective

The goal of this hackathon is to use data analysis and machine learning to predict customer interest. This means you'll determine identifying whether the customer would be interested in Vehicle insurance, based on past data.

## Hackathon site

https://www.analyticsvidhya.com/datahack/contest/janatahack-cross-sell-prediction
Note: You need to login to access the data
In case of trouble downloading the data, try using this link here to download

## Dataset

You'll be using the **Customer Dataset** taken from the hackathon site. The dataset contains **customer details along with their vehicle and policy information.** The dataset includes features like:

- Customer ID: Unique identifier for each customer
- Customer Demographics: Age, gender, region, etc.
- Vehicle Details: Vehicle age, damage, etc.
- Policy Details: Premium, sourcing channel, etc.
- Response: Target variable indicating whether the customer is interested in the policy or not.

## Step-by-Step Instructions

- Data understanding and observations
  Import Libraries, Load Data, Initial Exploration, etc.
- Exploratory Data Analysis (EDA)
  Visualize Distributions, Univariate Analysis, Bivariate Analysis, etc.
- Data Cleaning
  Check and Handling Missing values, Outlier Detection, Convert data types (if necessary), etc.
- Feature Engineering
  Encoding Categorical Data, Feature Scaling Numerical Data, Create New Features from existing columns, etc.
- Model Building
  Train-Test Split, Modeling using Pipelining, Ensemble Techniques, Model Training and Prediction.
- Model Evaluation and Tuning
  Evaluate Model Performance using Metrics specified(roc_auc_score) in the hackathon site, Optimize model by Hyperparameter Tuning using GridSearch or others, and Compare Models.

- Evaluate Test solution
  Use predict_proba (Probability of response 1) and test your solution in hackathon's website and capture result.
- Interpret Results
  Summarize your findings, model performance, and key insights into a final report. Compare models and explain which model were most impactful in predicting promotion and choose it as final model and pickle it for using them in API and UI.
- API Creation
  Create a FastAPI endpoint using the pickled final model.
- User Interface UI Development
  Create a Web App User Interface with the pickled final model using Streamlit library.
- Deploy API using GCP or AWS
  Create a new repository in your personal github account and push the required files (python files, Docker files, Requirements file, pickle file, etc.) and deploy them using GCP or AWS for public use and share the respective URL links.
- Deploy WebApp using GCP or AWS or Streamlit
  Create a new repository in your personal github account and push the required files (python files, Docker files, Requirements file, pickle file, etc.) and deploy them using GCP or AWS or Streamlit for public use and share the respective URL links.
- PowerPoint Presentation
  Prepare a PowerPoint presentation having all the steps performed along with summarizing the problem, approach, findings, recommendations and deployed links.
- Recorded Demo
  Record a brief demo (5 minutes) walking through your code, explaining your methodology, and showcasing results.

## Deliverables

**Files:** Submit all the Python code files, Docker files, Requirements files, Jupyter Notebook, Links of deployed API and WebApp(UI) link.

**PowerPoint Presentation:** Prepare a PowerPoint presentation having all the steps performed along with summarizing the problem, approach, findings, recommendations and deployed links, also convert them to a .pdf file as well

**Recorded Demo:** Record a brief demo (5 minutes) walking through your code, explaining your methodology, and showcasing results.

## Additional Notes

**Documentation**: Ensure your python code is well-documented with comments explaining each step.

**Submission:** Push all the deliverables in your github and share us your github link by posting thru Piazza or submit in this Google Form here

Good luck! We look forward to seeing your innovative solutions and insights. Remember, this hackathon is not just about reaching the answer, but also about experimenting and learning along the way. Happy coding!