Project Name – Sarcasm Detection

Deadline - 15 Days

Problem Statement –

This case requires trainees to develop a text classification model to label a news headline as sarcastic or not. This News Headlines dataset for Sarcasm Detection is collected from two news website. We collect real (and non-sarcastic) news headlines from different Post.

Expectations from the student:

1. Find the 6 topics related to news article headlines.
2. Do EDA to find the top 50 sarcastic words. Make a word cloud for top 200 frequent words.
3. Can you identify sarcastic sentences? Can you distinguish between fake news and legitimate news?

Data Set : Click on this link to download

Number of attributes:

Each record consists of three attributes:

1. is_sarcastic: 1 if the record is sarcastic otherwise 0
2. headline: the headline of the news article
3. article_link: link to the original news article. Useful in collecting supplementary data

Evaluation Basis

The project will be evaluated on the following basis:

1. The process of building the model should move from simple to complex. This means its mandatory to implement multiple linear regression and logistic regression before approaching any advance level algorithm such as Random forest etc.
2. Every model should be supported by reason of acceptance or rejection. Special emphasis on the reasons why the student has picked/dropped an algorithm.
3. The student should mention which error metric is used and why. For example, if RMSE used in place of MAPE, the reason should be clearly mentioned by the student that why a particular error metric is used. Special emphasis on goodness of fit such as AIC and confusion metrics.
4. The student should revise the concepts before starting the project work.
5. The student should be confident enough to explain every concept that is written in the project report.
6. The questions during the mock sessions would not limit to project. Any topic that is covered in the syllabus can be asked during the mock session.
7. If student unable to explain the project report or during the session, it appears that student has copied the solution then zero marks will be given to the student.
8. The code should be written keeping in mind the code file can be run from DOS prompt.
9. The instructions to run the code file should be submitted with the project report.

Deliverables from Candidate

1) Code written in Python.
2) Comprehensive Project Report.
3) Instruction to deploy and run code.

Always remember these evaluation basis, and your deadline. And your aim is to meet the deadline.

In case of queries, reach out to us at support@edwisor.com