Bidder challenge:

Background: A company acquires data called RTB (real-time bidding), which refers to the purchase of ads through real-time auctions. With real-time bidding, advertising buyers bid on an impression and, if the bid is won, the buyer's ad is instantly displayed on the publisher's site. For simplicity, let's assume the auctions are conducted as second-price auctions where the winner pays the price proposed by the second-highest bidder. Our goal is to find the optimal bid for each bidding opportunity. We do that by building a model that can predict the outcome of the impression based on different parameters.

A.
The purpose of this challenge is to propose the best bidding strategy based on the given data set. Then, use the proposed algorithm to bid on the data set provided in the validation set tab.

The purpose of this challenge is to propose the best bidding strategy based on the given data set. Then, use
the proposed algorithm to bid on the data set provided in the validation set tab.

Comments :
1. Need to pay attention that the training data can come in different scales and statistical calculations.
If a column or data still doesn't make sense, you are welcome to mention it and explain why.
2. Profit = Revenue – Cost

Input file: data.xlsx contains 2 sheets, with the following schema:
train_data tab: Aggregation of historical events

train_data tab: Aggregation of historical events

	Description
impressions	The number of impressions observed from this combination
clicks	The number of clicks observed from this combination
eDate	The date of the event
channel	The RTB channel we are bidding to
country	The country of the user
os	The operating system of the user's device (iOS / Android)
networkType	The network type to which the user's device is connected to (WIFI/ 3G)
deviceType	The device type (tablet/mobile)
publisherCategory	The category of the app we are presenting the ad in
advertiserCategory	The category of the impression we are advertising
product	The application Id we are presenting the ad in
subProduct	The app sub id we are presenting the ad in
campaign	The campaign we are presenting
advPackage	The package of the app we are presenting in the ad
advMaturity	The maturity of the advertiser (for adults/ kids etc.)
AverageWinPrice	The average price we actually paid for these ads
rate	The rate (payment) we get from the advertiser for each click we generate

*we= The company

Validation_data tab: Events to classify

	Description
sessionId	Unique identifier of the event
eDate	The date of the event
channel	The RTB channel we are bidding to
country	The country of the user
os	The os of the user's device
networkType	The network type to which the user's device is connected to (WIFI/ 3G)
deviceType	The device type (tablet/mobile)
publisherCategory	The category of the app we are presenting the ad in
advertiserCategory	The category of the impression we are advertising
product	The appId we are presenting the ad in
subProduct	The app we are presenting the ad in
campaign	The campaign we are presenting in the ad
advPackage	The app we are presenting in the ad
advMaturity	The maturity of the advertiser (for adults/ kids etc)
rate	The rate (payment) we get from the advertiser for each click we generate

Notes:
• Please provide the predicted CTR and proposed bid for every event in the validation set. Make sure the results make sense.
• It is perfectly fine not to fully complete the task. Please emphasize the results summary document. Explain your thoughts, what you research, and based on which assumptions you made decisions. Our first goal is to understand your way of thinking and only after that
check your coding abilities.

Part A Outputs: Please provide the output in the following format:
• A directory with the name "Assigment"
• Include a file inside the directory named bids.csv, containing the sessionId (of the Classify tab), and the CTR and bid predicted by the algorithm
• Code of the algorithm
• A clear results summary (pdf, ppt, doc) including your suggested bidding strategy and the means you used to evaluate your results. This file should also include your answer to questions B-C (see below)

Question B (no code):
In general, which bidding strategy would you suggest if the auctions are conducted as first-price auctions
where the highest bidder wins and pays the exact bid.

Question C (no code):
If we told you that the price the advertiser pays the company (rate) is estimated and isn't guaranteed. Meaning we know how much they will pay us not in real time but in retrospect. In real time we only have an estimated rate. How would that affect your chosen strategy? How would you

deal with the problem? Is it necessary to clean excessively high-rate values that probably won't be paid? How will you do it? Which algorithm will you use?
Why?

General concepts:

Impression: An impression is a metric used to quantify the number of digital views of an ad (when the ad renders on a user's screen, we count impression)

First / second price auction:

All bidders simultaneously submit bids so that no bidder knows the bid of any other participant

- **First-price auction -** the highest bidder pays **the price that was submitted**
- **Second-price auction -** the highest bidder pays **the price of the second-highest bidder.**