Capstone Project Proposal

Jorge Luiz Figueira da Silva Junior

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	The rise of Web 2.0 allowed the success of social networking platforms, and these in turn proved to be a very opportune way to sell products and services, as well as the strengthening of the company-consumer relationship. Social networks like Facebook help generate more sales with ads. One of the important elements of an ad is good copywriting, as it is responsible for creating product value and influencing the customer's decision and purchase process. Still in this panorama, researches claim that the future of marketing permeates artificial intelligence and that it will play at every stage of advertising development, evaluation and execution. In previous contact with some marketing agencies, the potential use of data obtained from the Facebook Ad Manager tool was identified. Analyzing this data through ML/AI can identify variables that drive more effective advertising and predict improved performance. In addition, they can help with the ad creation process and test ad impact (Moses et al. 2021). In this sense, the business objective is to produce an AI-based tool that can be used as support in the process of creating good copywriting for ads on social networks. Moses, E., Clark, K. R., & Jacknis, N. J. (2021). The Future of Advertising: Influencing and Predicting Response Through Artificial Intelligence, Machine Learning, and Neuroscience. In Handbook of Research

	on Applied Data Science and Artificial Intelligence in Business and Industry (pp. 151-166). IGI Global.
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	This is important, as it not only improves the quality of an ad's persuasion (through pre-identification by the model), it also reduces the money used in the campaign, as less running time may be needed for the ad to reach a good performance. This project is proposed to generate revenue for our team as follows: Private Partnership. Given that the dataset required for the AI product belongs to the interested parties, it is believed that the best way to monetize the product is through private partnership with marketing companies that want this feature incorporated into their ad creation processes.
Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	The proposed project will use ML/AI to predict potential copywriting used for advertisements. For this, the models used neural networks to automatically predict whether a copywriting is “Weakly Performable” or “Strongly Performable”. As direct results there is an increase in performance and improvement in the quality of copywriting of ads. As indirect results, a reduction in the amount of money used to serve the ads is expected.

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	The main metric of success will be compared the performance of the model based on real world data. Periodically, project partner marketing specialists would make the data available from the Facebook Ad Manager, so that they can monitor whether the model is hitting the predictions well.

Data

Data Acquisition Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	The data will be obtained through the Facebook Ads Manager tool of the marketing companies that are partners in the project. The data exported from this tool is tabulated data (CSV file) that contains useful information such as the start and end date of the ad serving, ad text (copywriting), number of reactions (Like, Love, Care, Haha, Wow , Sad or Angry), number of shares, comments and calls to action (number of clicks on the link). The collected data do not have any elements that can characterize or identify a user of the social network, thus, they are not configured as sensitive data. There is no additional cost involved to obtain this data, and it can be made available on an ongoing basis via the Facebook tool.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	The tool proposed in this project will be built in a multidisciplinary team, where marketing professionals acted strongly in the development and validation process since the beginning of the project, thus ensuring an improvement of data and reduction of anomalies. After data collection, an exploratory analysis will be performed. This will define arbitrary values of quantities of reactions, comments and calls to action, so that ad data can be divided into Class 0 – “Weakly Performable” and Class 1 – “Strongly Performable”. In this sense, the labeling work will not be done subjectively, but objectively based on the criteria listed by marketing specialists. The ad data collected may be from different market niche types, geographic domains, and audience types. This can influence the performance of the model. Therefore, among the biases incorporated can be exclusion bias and association bias.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels versus any other option?	As initially the interest of the problem is to know whether a copywriting is potential or not for an ad, it is configured as a problem of binary classification, that is, we will have Class 0 as labels for " Weakly Performable" ads and Class 1 for "Ads" Strongly Performable”. Keep in mind that in the future a label will be included to describe an

	intermediate category.

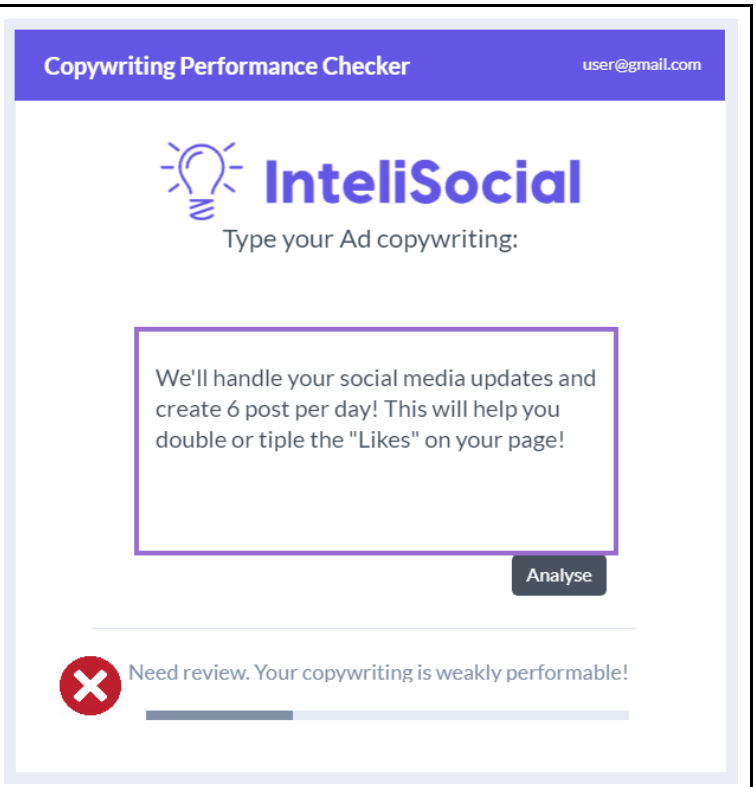
Model

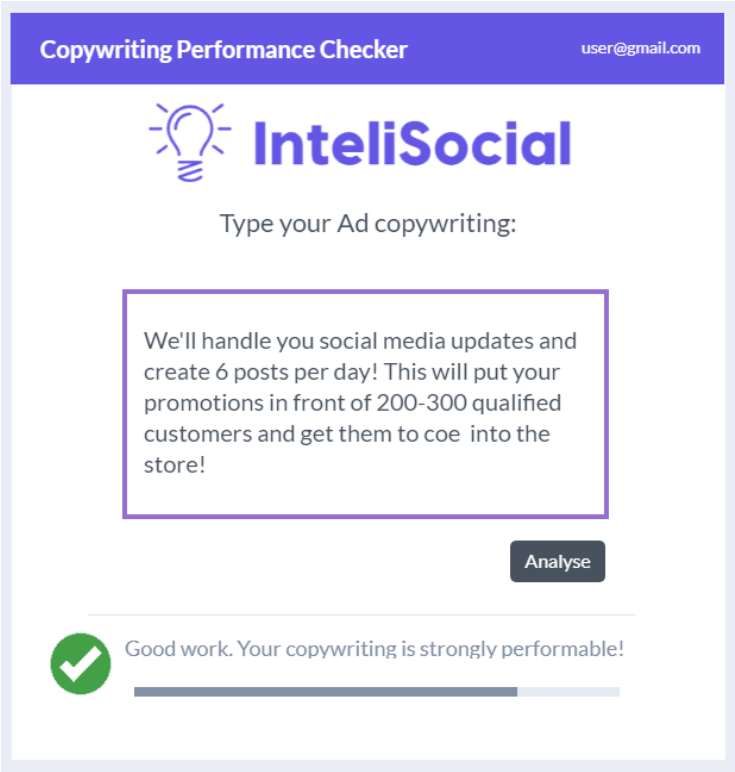
Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	In this project, an advantage of using an internal team for model training was seen. While the team's main objective is to train the prediction model of potential copywriting for ads; having an in-house team is a great advantage to produce adjacent studies and analyses, for example, ML techniques can be useful to identify patterns, extract knowledge and help marketers identify which elements are more assertive to produce a potential ad. In addition, the advantage of an internal team is being close to marketing professionals, who in this project also act as stakeholders, this helps to ensure a result more in line with the expectations of the partner companies in the project.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	The metrics that can be used to evaluate the results are the classic ones reported in the ML/AI literature, such as ROC and AUC curve analysis, precision and recall. A performance level of at least 85% was defined, as this is a product that supports financial decision-making (deciding whether a particular ad is served or not.) Here we always want to choose the copywriting that will lead to the best performances for that we need the lowest possible rate of False Positives and False Negatives.

Minimum Viable Product (MVP)

Design

What does your minimum viable product look like? Include sketches of your product.



	
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?	The proposed tool, here called “InteliSocial” is intended for marketing agencies, media companies, marketers, advertisers, influencers and freelancers. Main case: a client company of a marketing agency needs to start a campaign in order to attract potential customers to their website. However, this company has budget restrictions that prevent it from investing in a longer campaign airing period. Marketing strategists at this agency are challenged to make an ad that performs well. The product predicts whether the text used in the copywriting of the ad is “Strongly Performable” and aids marketers in the creative and decision-making process. Access to this product is made during the private partnership between marketing companies interested in adopting this technology and the project team. During the partnership, a schedule is made for the implementation of this product in the companies' process and operation.
Roll-out	As mentioned previously, the development of this product would be done via private partnership with

How will this be adopted? What does the go-to-market plan look like?	interested marketing companies. There is no need to develop a market plan, hoping that the successful results obtained from a partnership will create a spontaneous demand. In addition, the companies previously contacted for the development of this project are part of franchise networks, that is, there is a significant potential for expansion. The launch of this product can be organized and divided into major milestones, namely: 1) Hire the team A good team would consist of: 2 Data Scientists, 1 Machine Learning Engineer, 1 Software Engineer, 1 Quality Assurance Analyst and at least 2 marketing professionals (partner companies). 2) Analyze and extract the dataset The datasets will be built using the data obtained from the Facebook Ads Manager of the partner companies in the project. There is no need to assemble an extra team to label the data, as the labeling will be done automatically and objectively following criteria listed by marketing experts. 3) Build and train the model An initial model is expected to be developed over 2-4 weeks. 4) Test model and performance analysis The developed model will be tested, its performance evaluated, and if necessary, a refinement will be carried out. It is expected that 2 weeks will be invested for this task. 5) Launch After the model has reached the minimum performance expected by the project team, the development of other parts of the product will be done, such as the interface and integration of other modules for use by the end user. The launch is believed to be done in 3 months. 6) Further improvements

	We plan in a next release to inject a new dataset into the model, this will be done periodically or when performance drops over time. In addition, in this current version, only the copywriting part of the ad is being analyzed, which is mostly composed of text, and the use of hashtags and emojis. Believes that the consideration of the creative of the ad (video or photo) can be used as part of the dataset and thus increasing the classified performance of the model, making it possible to predict a potential ad, considering all its elements.

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	The project's partner marketing specialists will continually make Facebook Ad Manager reports available. From there, the predictions will be confirmed and serve as feedback to the model. Over time, more labels can be added, and the model retrained on these updated labels. Subsequently being tested to ensure efficiency. The data used in training does not differ from real world data. Because the data used during training on the model is real world data curated by the project team. After the addition of the new model, 20% would be used on the updated model and 80% on the current model, and performance compared.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	Periodically, the project's marketing experts will perform a performance analysis of real-world data and compare it to the results predicted by the model. Feedback obtained by experts would be used to mitigate unwanted bias.