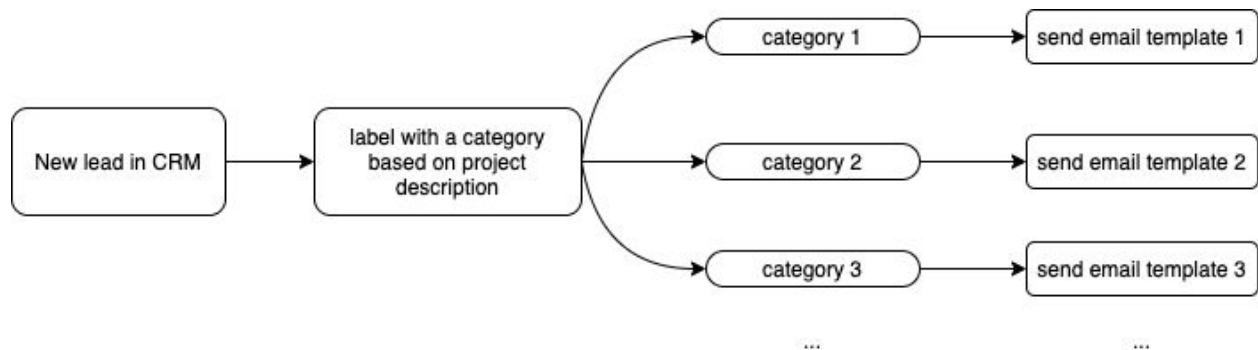


Send custom emails to leads with Machine Learning

Capstone proposal - Machine Learning Engineer Nanodegree

Domain background

I am the founder of a digital agency called DeepIdea Lab. The services we offer are website creation, digital marketing, and graphic design. We have a partner who is sending us leads. The lead details include a description of the project each lead is looking for. This data is automatically pushed into our CRM. The next step is to send an email to the lead using one of our templates depending on the project's description. However, we would like to automate this part. This is the topic of this capstone project: using machine learning to send a custom email to the lead based on the project's description.



Problem statement

Sending a manually composed email to every lead is time-consuming. We already have a few email templates available to be sent. However, so far, someone needs to read the project description in order to choose the right template. This is a typical natural language processing problem.

Datasets and inputs

The dataset available includes descriptions of around 100 projects. A corresponding category needs to be attributed to every project description. For privacy reasons, the dataset also needs to be cleaned of the personal information (name, website, email, phone number). This will also improve the training of the model.

Solution statement

The solution is to create an algorithm that will attribute a category based on the project description. As soon as a new lead arrives in the CRM the algorithm runs and classifies the lead into a category and sends the corresponding email.

Benchmark model

The sentiment analysis model deployed in one of the last projects can be a good benchmark and can also serve as an example as it also tackles an NLP question.

Evaluation metrics

Based on the dataset we have, a decent evaluation metrics would be 80% accuracy on validation data. A confidence score can also be considered as evaluation metrics.

Project design

A good part of the project consists of cleaning and anonymizing the data. The cleaned project description should contain only lowercase letters, etc. We also need to decide how many and which categories we employ. An example of a project category would be “website”. When the dataset is clean and labeled we can train an algorithm to predict the category.

