

App: "Predicate"

What is the purpose of the app or product? What are you trying to accomplish?

The purpose of the app is to provide a web platform, where users can compete for predicting properties of data sets.

What is the current problem(s) it will solve?

Companies frequently have data but not enough resources to analyze it. Instead of hiring an analyst, a company with a data analytics problem could work with the app administrators who will upload a data set to the web app. They will turn the data prediction problem into a prediction competition (with a deadline and optionally a financial incentive for winners).

How will it streamline or improve the current process or facilitate a new process?

The closest existing solution is <https://www.kaggle.com/>. By having data scientists compete for challenges and prizes, they facilitated efficient solutions to difficult problems.

I will be teaching a machine learning course next spring. I want my students to be able to sign in into the web app, and complete homework in form of prediction challenges. I am not aware of a similar solution that would streamline this.

What is the product vision?

I have the vision that instructors of courses related to machine learning can use this app to facilitate competitions among their students in terms of who creates the most accurate machine learning models and data predictions. If successful, the scope of the app could be increased later to engage businesses to submit data and data scientists to participate in analyzing data.

What aspects are already in place? What needs to be added? Will the app need to be started from scratch or can you leverage existing assets?

The app needs to be started complete from scratch. I could imagine if a prototype works well, that I would hire a follow-up gig with you for additional functionality.

What should the app be able to do (ie. functionality)?

Here is a list of "classes" I could imagine:

User:

Can register by providing Firstname, Lastname, username, email and password

Each user has a unique username and unique email address

Each user can participate in zero, one or several challenges (see below)

Have user landing page where they see in which challenges they are participating in and what predictions they submitted

Users can only view their own user landing page and not the landing page of other users (unless they are administrator)

Dataset

A Dataset has at least 4 elements:

- A text field with a unique name
- A text field with the content
- A text field with a description
- A reference to the Challenge (see below) it belongs to

There is no need to have a separate listing of all data sets. Instead, datasets should be listed as part of “Challenges” (see below).

Challenge

A Challenge correspond to machine learning prediction challenge.

A Challenge contains at least 4 elements

- A text field with a unique name
- A text field with a description.
- A date corresponding to the deadline
- Possibly several Dataset entities which belong to a Challenge

For each challenge, there is a web site with a description of the challenge.

Each challenge has zero, one or several dataset s that are available to registered users for download

Each challenge has a potential deadline

Each challenge has a leaderboard (sorted list) with the latest prediction (see below) of each challenge and how they are scoring with each challenge

The administrator can manually add, remove and edit challenges. The administrator can edit the description of the challenge and upload/edit/remove dataset files.

Prediction

A prediction contains at least 5 data elements:

- A text field containing the actual prediction provided by the user. This web app does not have to be able to further parse this text field
- a data of creation (to be created automatically)
- A reference to the User who created it
- A reference to the Challenge to which it applies
- A score. Users can view the score, but not edit it. Only the administrator can edit the score

A prediction is created by a user with respect to one Challenge

A prediction contains a text section that the user provides upon creation of a prediction.

Users cannot remove predictions.

A prediction contains a score. Initially, this score will be N/A or a number between 0 and 100. The administrator can manually set the score of a prediction. In later versions of the web application, the score may be automatically generated by the web app as a function of the upload prediction text.

A user can have 0, one or several predictions with respect to one or several challenges

Features Are?

Functionalities:

- Users can register for the web app
- Users can view their landing page: on there, they can see what challenges they are signed up for and how well they are doing in them.
- Users can browse a list of challenges.
- For each challenge there should be a web site that shows a description, deadline and data sets. Also shown is for each challenge a leaderboard of scores of user submissions.
- Users can sign up for or sign out of challenges.
- Users can download datasets that correspond to a challenge for those challenges they are signed up for.
- Users can upload text data for each challenge.
- The administrator of the app will evaluate the text submissions manually and score them.
- The submission scores entered by the administrator should lead to updated leaderboard for that challenge