



ONSTRUCTING INTELLIGENCE

Mentors - Michael Fairchild
Carlos Torres
Ramsay Stirling

Procore Technologies - Machine Learning
Capstone 2016-2017
Vivek Patel - Team Lead
Nicole Moghaddas - Team Scribe
Lauren Dumapias, Deanna Pham, Charlie Getzen

>> Motivation <<

Two main challenges of Construction Projects:

- Staying on schedule
- Safety

These can delay the finished product and can be PRICEY!

>> The Data <<

Submittals > New Submittal

New Submittal

General

GENERAL INFORMATION

Title:

Spec Section: Submittal Package:

Number & Revision: - Status:

Responsible Contractor: Received From:

Submit By: Issue Date:

Received Date: Final Due Date: No approver due dates have been set.

Sub Job: Cost Center:

Submittal Manager: Type:

Private: ☒ Visible only to admin, workflow, and distribution list members. Location:

Description:

Submittals are required primarily for the architect and engineer to verify that the correct products and quantities will be installed on the project in compliance with the design documents/contract documents.

Observations > New Observation

New Observation

Type: Status:

Title:

#: Priority:

Location: Trade:

Assignee:

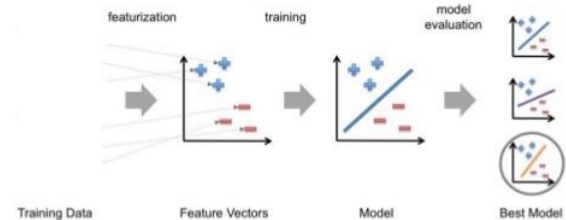
Due Date: Private: ☐

Description:

Attachments:

Cancel Create

Typical Steps in ML Pipeline



An observation is an indication of a hazard or defect on a project, made by any worker on the jobsite, usually the foreman.

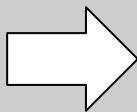
>> Features & Targets <<

Submittals

cached_distribution_sent_date → the issue date of a submittal

due_date → the date a submittal should be completed by

submittal_type → the category of a submittal (i.e. materials, plans, etc.)



is_late → binary indicating if submittal is late or not

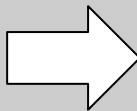
num_days_late → how many days late a submittal is

Observations

attachments_count → how many media attachments an observation has (could potentially indicate urgency)

status → what stage the submittal is in

priority → indicates the importance of the observation



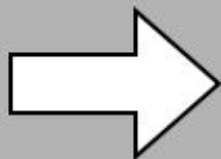
observed_safety_score → cumulative score based on number of attachments and urgency/status of observation

>> Our Solution <<

- Implement machine learning techniques
 - Train models using previous company data
 - Create a company's “schedule” and “safety” score with these trained models
-

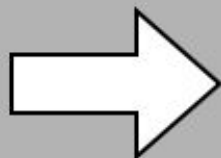
Dummy Coding

Color
Red
Blue
Orange
Red



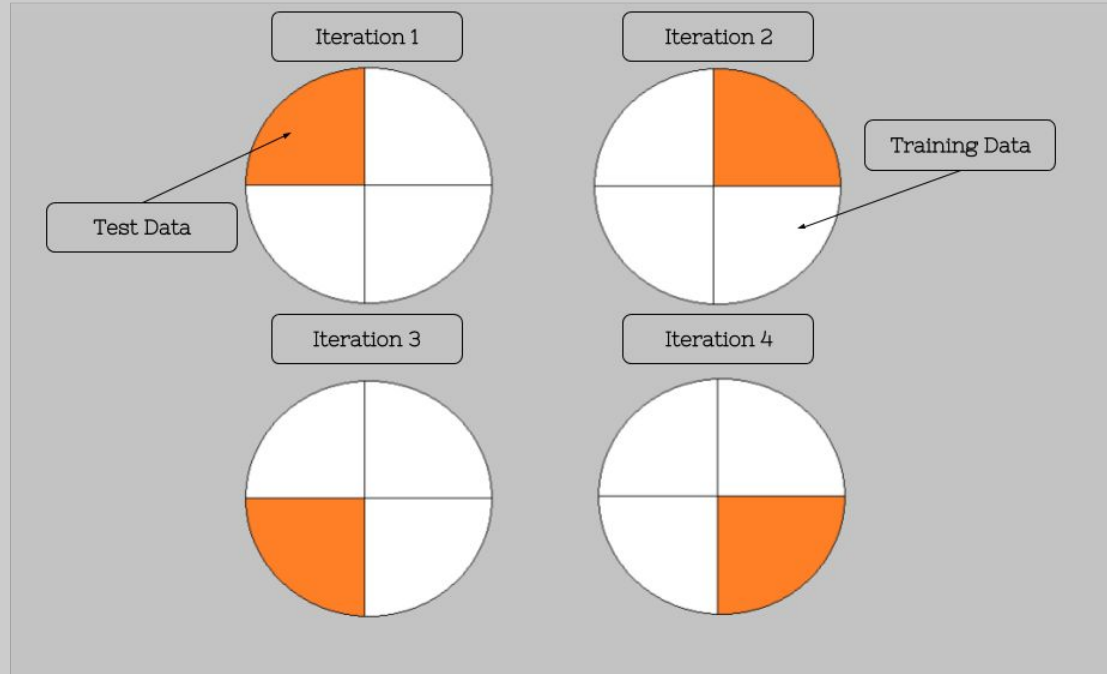
Red	Blue	Orange
1	0	0
0	1	0
0	0	1
1	0	0

Categorical variables



Dichotomous variables

>> K-Fold Cross Validation <<



*Running on 4 folds ($k=4$)

DEMO

>> Architecture <<

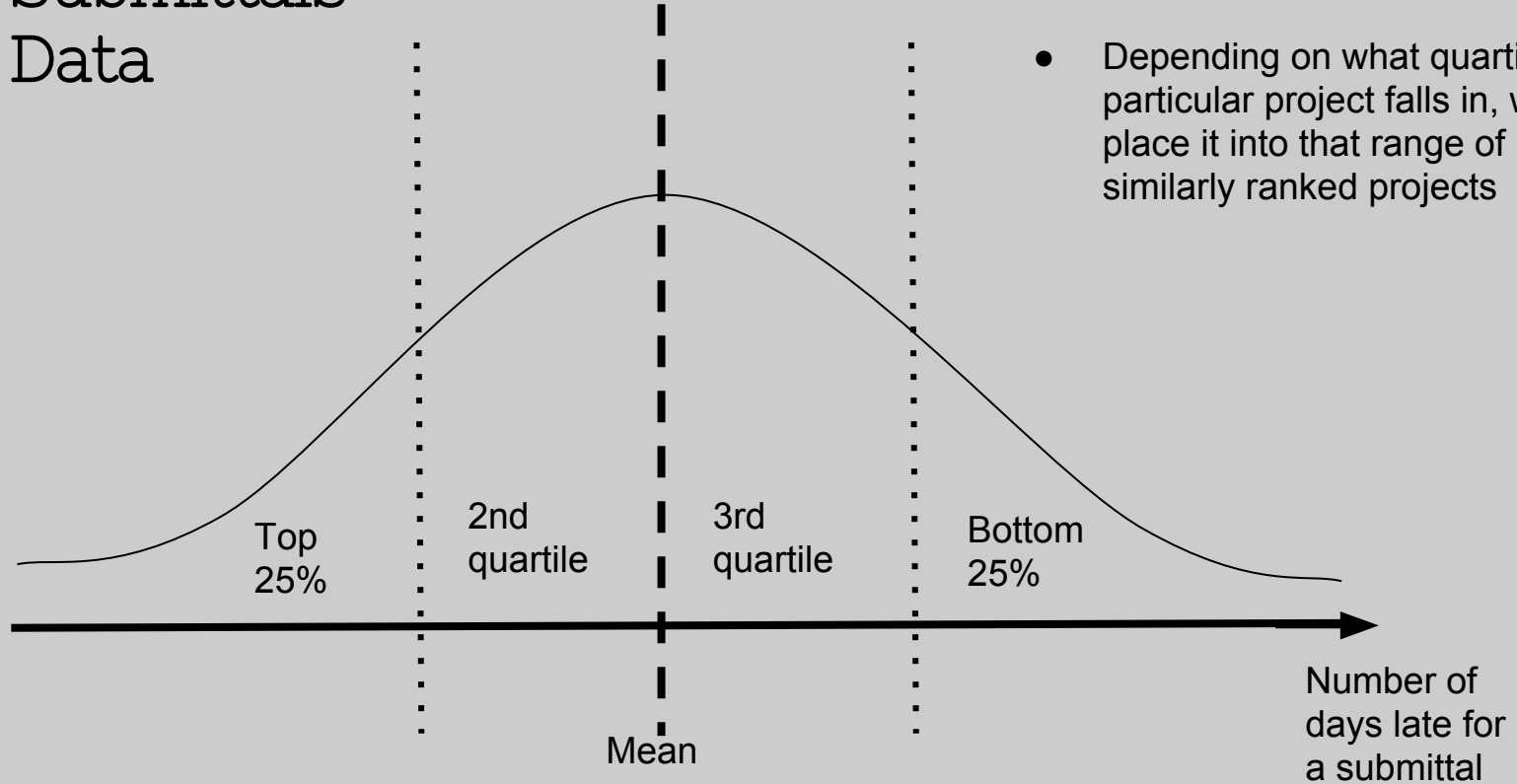


Inside AWS instance:

- I. Model is loaded
- II. Project data requested
- III. Data is preprocessed
- IV. Data is run through model
- V. Score is calculated
- VI. Score is send back to extension

Submittals Data

Number of Submittals





Your **schedule** percentile: **Top 25%**

Improvements:

- Increase submittal_type

Your **safety** percentile: **Bottom 25%**

Improvements:

- Increase length_of_description
- Decrease attachment_count



Your **schedule** percentile: **Top 25%**

Improvements:

- Increase submittal_type

Your **safety** percentile: **Top 50%**

Improvements:

- Decrease attachment_count

>> The Future <<

- Understand project progress and safety based on statistical evidence
 - Lay the groundwork for future predictive analytics (i.e. financial data)
 - Potential to make predictions for projects outside the construction industry
-

Thank you!

(mic drop.)
