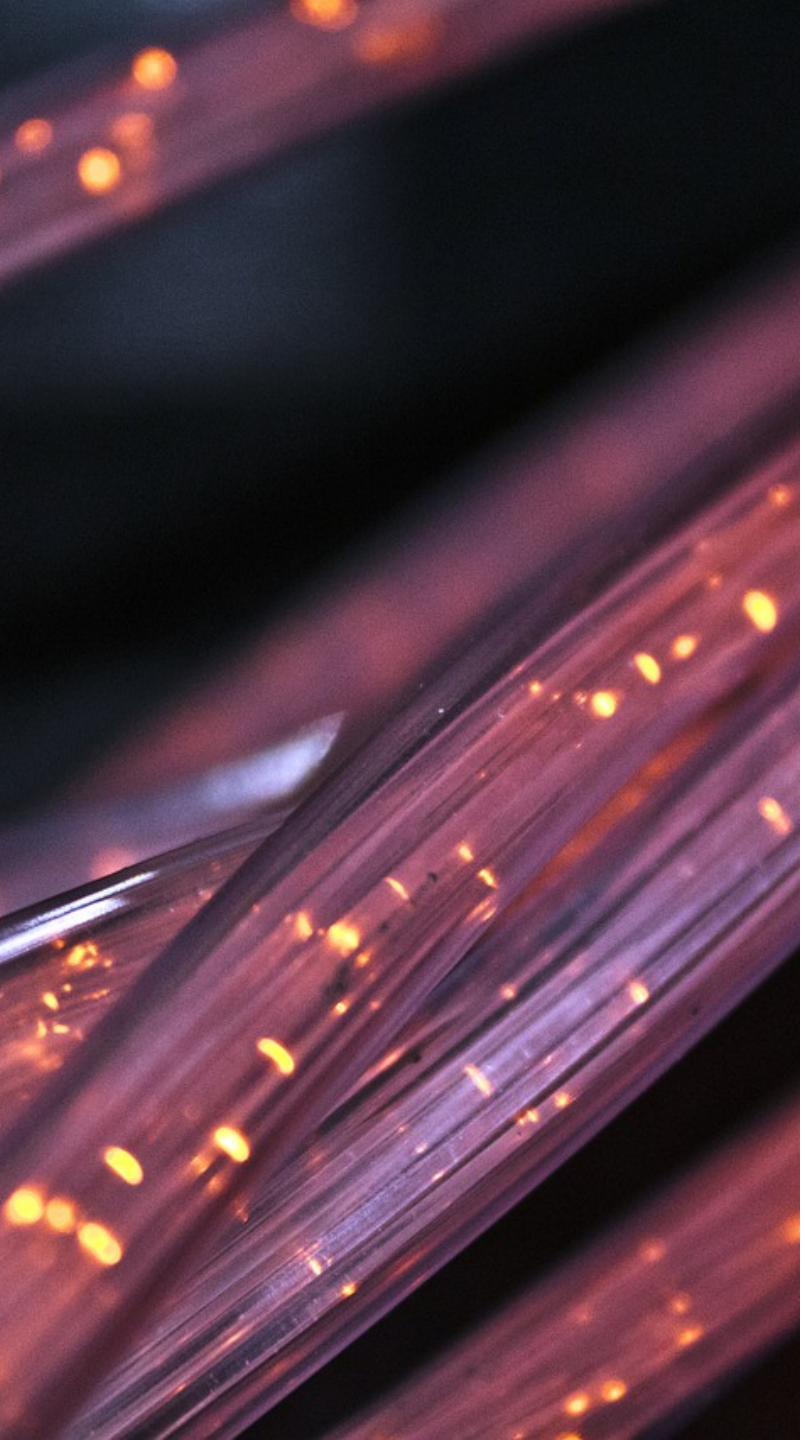


Deep Learning Alpha

Project Proposal

Sergio Rio & Jason Wheeler

MADE WITH beautiful.ai

Deeplearning.ai's Mission

- Deeplearning.ai seeks to make a world-class AI education accessible to people around the globe so that we can all benefit from an AI-powered future.
- Deeplearning.ai's mission is our mission.

MADE WITH
beautiful.ai

Project Summary

- Deeplearning.ai develops courses on key applications of AI.

- Their database includes approximately 300k individuals.

  Each individual's record has activity like opening marketing emails, viewing web pages, and registering for events. Some also have data on their location and job title, and newsletter subscriptions: The Batch, Events and Course Announcements.

- Deeplearning wants to understand the journey a potential student takes from initial registration to course completion.
  - **A user's journey can include:**
  - Page views
  - Interaction with marketing emails across mailing lists mailing lists
  - Registration for local community meetups, expert panels and learner community events events

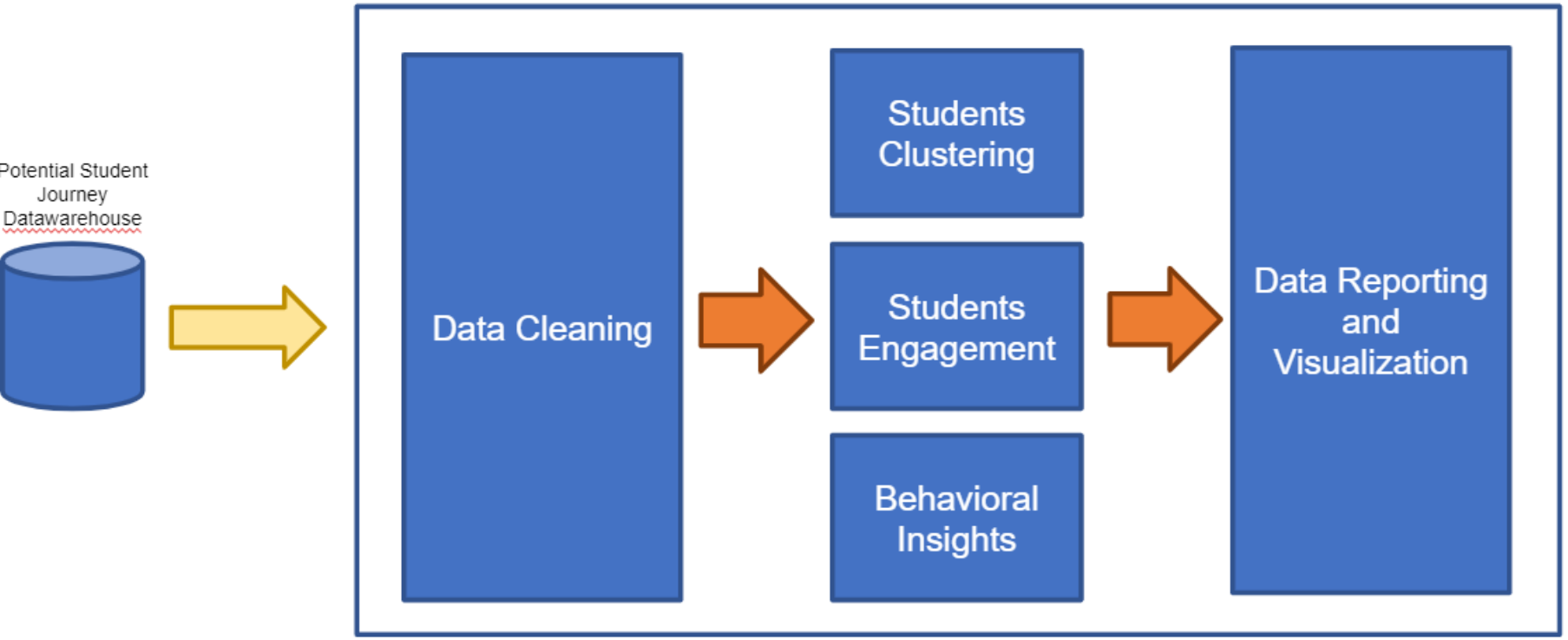
MADE WITH beautiful.ai

Scope of Work

- Understanding the problem and potential solutions
  - Data sets analysis
  - Create Dictionary
  - Identify and close any information gaps
  - Read relevant research papers
  - Develop initial product design
  - Define benchmarking baseline
- Develop product prototype
  - Build Data Collection
  - Build Data Cleaning Module
  - Build Feature Analysis and Selection Module
  - Build Machine Learning Models
  - Build Reporting Module
  - Test and validate prototype baseline performance
- Build, Test and Run Product Version 1 Pipeline
  - Build Product Version 1
  - Test Version with fresh data
  - Production deployment
  - Measure Product Version 1 performance
- Product presentation and delivery
  - Production overview and training
  - Product acceptance (User Acceptance Testing)
  - Project wrap-up
  - Final project presentation

MADE WITH beautiful.ai

DeepLearning.AI Insights – System High Level Design



1. Journey map of general population and paying learners
   a. What is the usual flow through our pages, events, emails to paying for a course?
   b. What are the clusters of users?

Engagement models to increase course registrations
   b. which groups are the most likely to register with X engagement

2. Any additional insights about our paying learners if available

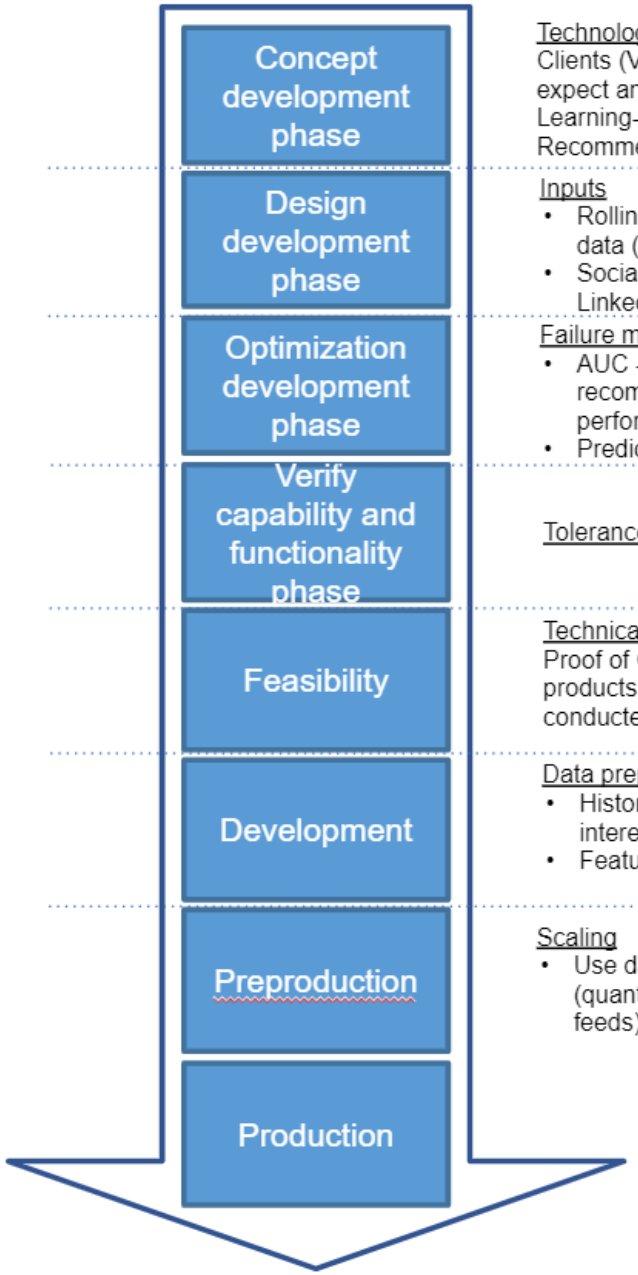
MADE WITH beautiful.ai

Appendix

MADE WITH
beautiful.ai

Insights - System Life Cycle End-2-End Product Design

Design for Six Sigma: Roadmap for Successful Corporate Goals, Salman Taghizadegan, in Essentials of Lean Six Sigma, 2006



Technology Concepts
Clients (VPs of Sales and Operations) expect an innovative Machine Learning-based New Product Recommendation System

Goals
Maximize the return over the investment and revenue growth from new products

Constraints
- New system must be released to production in 3 weeks
- 3 years of past customer journey history
- 1 Data Scientist (part-time) and 1 ML Engineer (full time)
- Dedicated computing environment with GPUs/TPUs

Inputs
- Rolling window of historic customer journey transactional data (9 months for training and 3 months for validation)
- Social media feeds from Facebook, Instagram and LinkedIn

Outputs
- Recommendation of products to be launched (0: not recommended 1: recommended

Flow Diagram(*)

Data Privacy and Bias
Customer and PII will be removed before used by model data pipeline

Failure mode effects analysis (FMEA)
- AUC - ROC curve to measure recommendation (classification performance)
- Prediction accuracy/recall/f1-score

Random baseline: run random classifications on historical data
Human baseline: gather performance of existing R&D and Marketing teams on new products recommendations
Simple heuristic: gather voice of the customer and customer surveys

Tolerance design analysis

Technical Feasibility
Proof of Concept with top 100 products by sales volume will be conducted

Economic and Operation Feasibility
confirmed by project budget and available company resources and specialized personnel

Data preprocessing & representation
- Historical data will be pre-processed to extract features of interest
- Features will be analyzed using Random Forest and OLS

Model Selection
1. Iterative process to select best model:
2. TPOT(Auto ML) for classifier based on quantitative historical data
3. Co-training for classifier based on social media feeds

Training & Hyperparameter Tuning

Feature Engineering

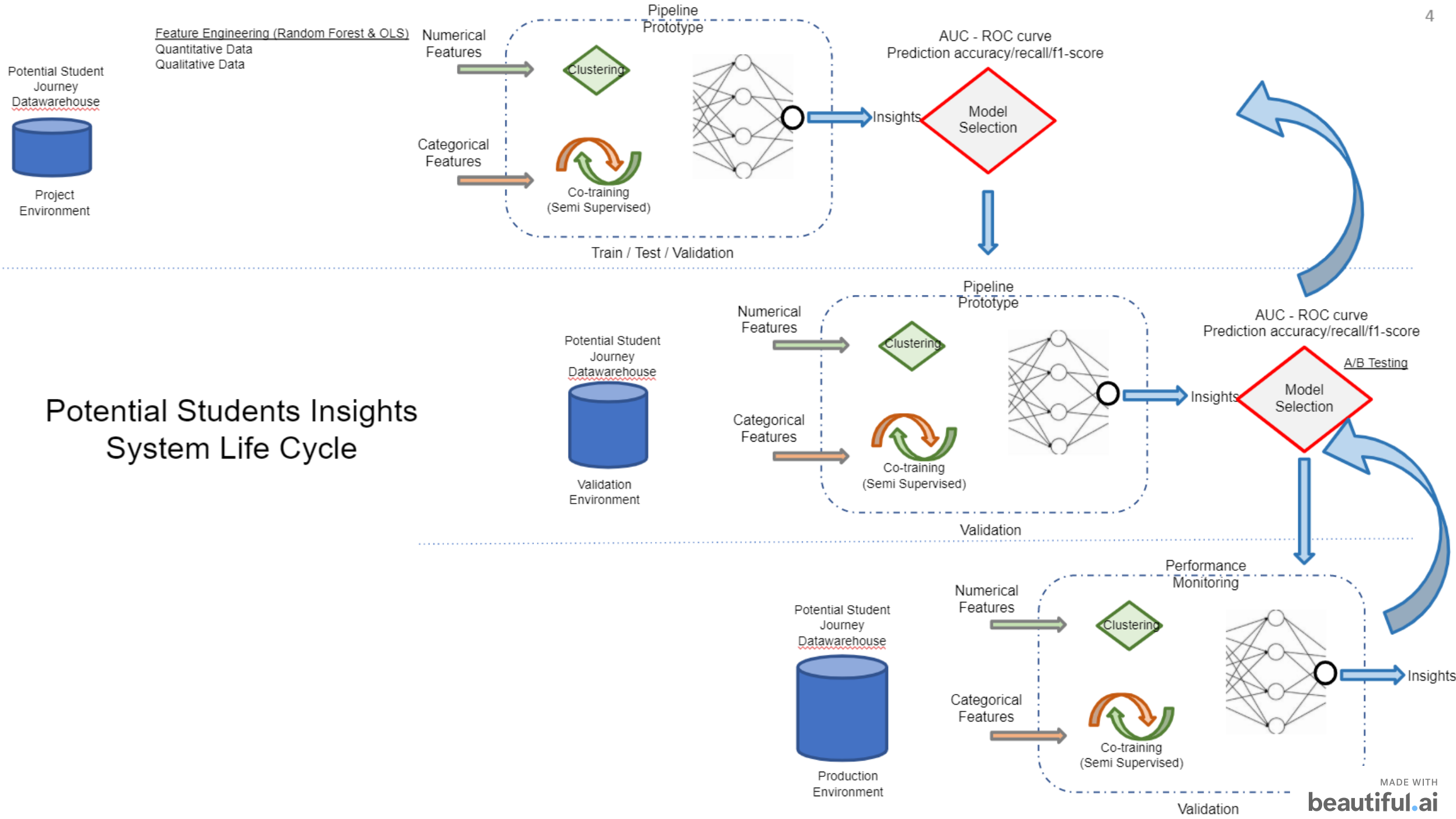
A/B Testing

Scaling
- Use data parallelism for training model pipeline (quantitative data) and co-training (social media feeds)

Serving
- Run experiments with Sales and Operations planners to confirm end user experience, predictor behavior, boundaries, errors, results interpretability and model explainability
- Conduct ablation studies -- removing each component while keeping the rest -- to determine the efficiency of each component. Find components whose removals eventually don't significantly reduce the model's performance but significantly reduce its complexity.
- Check model for propagating any gender and racial biases (if only if gender and race are features used)

MADE WITH beautiful.ai

Potential Students Insights System Life Cycle



MADE WITH beautiful.ai

Q&A

MADE WITH
beautiful.ai