SE 491 WEEKLY REPORT 4 - Tech Research Spikes

sdmay20-25: Consumer Aware Warehouse Management

10/15/19 – 10/22/19

Team Member	Roles
Jimmy Paul jpaul@craftydelivers.com	Client
Goce Trajcevski gocet25@iastate.edu	Advisor
Lindsey Sleeth lssleeth@iastate.edu	Meeting Scribe Project Manager Software Developer
Sam Stifter stifter@iastate.edu	Test Engineer Software Architect Software Developer
Omair Ijaz oijaz@iastate.edu	Quality Assurance Engineer Meeting Facilitator Software Developer
Jameel Kelley jamkelley22@gmail.com	Report Manager Software Architect Software Developer
Andrew Smith arsmith3@iastate.edu	Database Administrator Quality Assurance Engineer Software Developer
Elijah Buscho elijah@iastate.edu	Test Engineer Software Dev Proj Manager

Weekly Summary

Objective

The objective of the week was to set up a framework for the backend, as well as, set up a framework for the front end. Individually, each team member must needed to research product forecasting and document it in our wiki.

Tasks Accomplished

- Revision of 1st version of design document with advisor
- Individual research on methods for product forecasting
- Backend infrastructure setup
- Frontend infrastructure set up

Summary of Weekly Advisor Meeting

- Think of the context diagram as a workflow
- Pick one type of a data input and run the entire workflow and scenario with it and see the interaction among the main components. Use those interactions to generate the architecture diagram and system diagram.
- How does this affect the architecture? We may need to put several tables together or have flexibility of the attributes in the table.
- Put some instances back to back. We know how each step triggers in a scenario
- The moment you extend a scenario, you have to cater to having multiple data warehouses
- Be agile in the requirements, design, and the development process. Pick something and follow consistently and see where the questions start popping up.
- Can you make any inferences (normalizing) across data?
- We are doing a reactive type of learning because of the semi-agile manner that we need to work in.

- The moment that you start implementing changes in one area of your workflow, it will tell you radical changes in your workflow, but also help you identify testing scenarios.
- There will come a point when we might want to abandon a solution and a point when it would become too costly to abandon, but we will figure it out.
- Say BECAUSE of a,b,and c we can make a justification for our solutions
- We have enough background knowledge so we can try and unify it in writing. Start, add-on, abandon or finalize
- Make notes about discoveries and disagreements
- The moment you figure this out you can make justified decisions
- We might notice testing patterns and details that may arise in other scenarios
- Ask him to pick one customer and 3 products
- OR better yet, we can pick our own and ground it
- It is much easier to generalize from a ground example

Second Aspect to Think About

- Develop an application for December that is going to illustrate a tiny little bit of a use case scenario and show that alongside our presentation
- Say by the way, this is what we expect a casual user to have
- When we start focusing on the development in the next semester, this will be implemented here and like this and this is one example of a testing plan that we would like to have
- In week 15 we are presenting

What is the presentation is going to look like?

- We will have roughly 25 minutes for a presentation
- We will each get to babble a few words about our slides and there must be a good connection and flow between them
- Ground and expand our development ideas to help structure this development
- Refer backwards to things that we have done
- This is our closure, but here are possible extensions
- Behaviorally for the presentation there are 2 extreme points, we should not be hesitant about explaining
- There is a large document and we should not even try to cover all of the points in it. Make sure that everyone gets the big picture and everyone that listens to us identifies what are the main components and why they are there. We can always refer them to the document because they are supposed to read from the document

- Practice a little bit and make some eye contact during the presentation
- Repeat questions asked in the middle of the presentation

Past Week Accomplishments (Individual)

Lindsey Sleeth

This week I worked on technology research spikes. Specifically, I researched what the best languages are for product forecasting, which I found to be R and Python. Our team suspected and had done some research earlier in the project to know that we would decide between R and Python for our solution, however, we still were undecided on which technology we would use. The team was swaying towards using R because product forecasting relies very heavily on statistical modeling and R was written for staticians by staticians and has some very powerful modeling libraries. After some more research and discussion, we decided to use Python and the pandas library because most of our team is familiar with Python and for the scope of our will accomplish the same thing in R, but we will have the advantage of knowing Python. Additionally, I documented several articles in our wiki that will help us determine which method to go with for the product forecasting.

Among the methods that I researched, these were the main ones that I discovered:

1. Naive forecasting - comparing past data of a given time period and using it to predict future results
2. Demand forecasting - taking outside variables (weather, seasonality, etc.) into account
   a. Qualitative Techniques - relies on having sufficient data, past data + math = future performance, often looks at outside influences (product launches, holiday season, all available past data). Listed are some qualitative methods that require more research.
      i. Exponential Smoothing
      ii. Box-Jenkins Model
      iii. Time Series Model
      iv. Seasonal Indexes
      v. Causal Models
      vi. Moving Averages
      vii. Data Mining
   b. Quantitative Techniques - subjective viewpoints of market plays, relies on educated guesses based on experience and knowledge of how the market

works. Should not be relied on by its own, but if used with quantitative techniques, then we can be much more precise.

    i.   Delphi Method
    ii.   Prediction Method
    iii.   Historical Life-Cycle Analogy
    iv.   Game Theory
    v.   Executive Opinions
    vi.   Consumer Surveys
    vii. Sales Force Composite

Each of these areas requires more research. Because our client has so much data available, we will likely use the Client's current solution as a baseline for performance. It is slightly better than Naive forecasting currently. We will likely prototype a solution using a qualitative method and then incorporate a quantitative technique if time allows.

Elijah Buscho

This week I learned a bit about deep learning neural networks. A lot of research still has to be done into this area, but I came up with a basic idea for a potential solution:
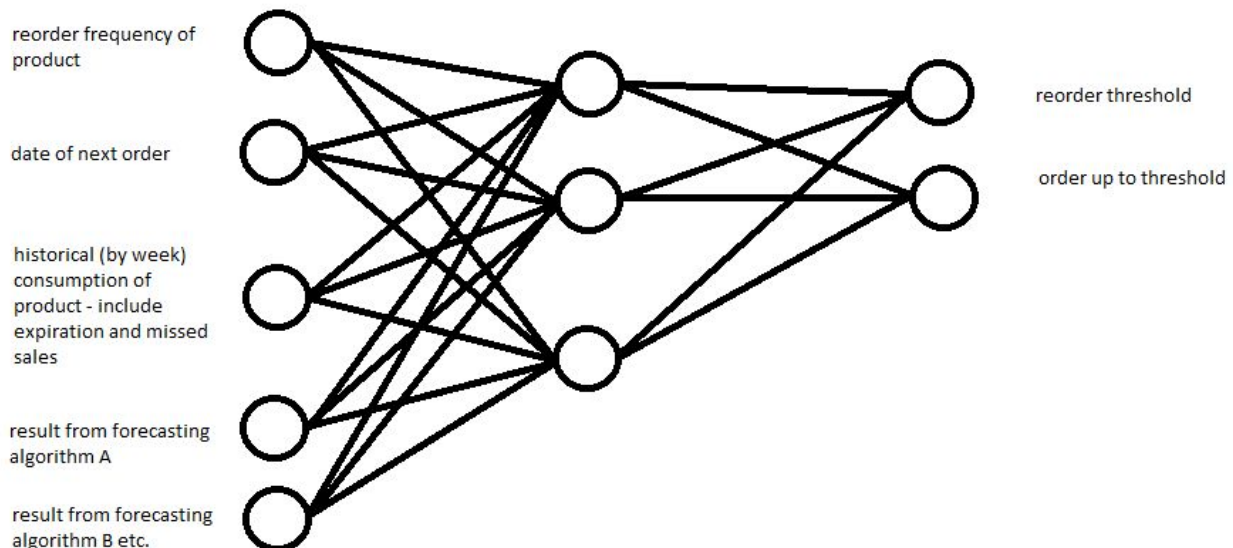
This diagram shows the neural network for a specific product. This will be expanded to consider all products that Crafty sells. The number of nodes in the hidden layers and the number of hidden layers is completely arbitrary at this point. A discussions with our client lead me to making the output of the algorithm to be the thresholds that are used in

their current system. Basically, a neural network solution is a solution that a human brain would discover if all it did was focus on the one problem it is solving, and you want to give it as much data to consider (inputs) as you can so that it can come up with the best solution. A human trying to solve this problem of prediction has already come up with a few algorithms to solve this (the forecasting algorithms that Lindsey researched), so in order to save on processing power so to speak of the neural net, we can just feed the outputs of the algorithms directly into the neural network.

Jameel Kelley

This week I centered most of my focus around doing Technology Research spikes. This included pipeline research for expedited ci/cd, frontend testing research to determine the best framework to use, and analysis of typescript to determine if it will be desirable to use.

CI/CD was made easier as we had previously done something similar with a group project with both a frontend and backend. However, this was not perfect and still needed work to fix. We are still working out some of the kinks to ensure the pipeline will be robust.

Frontend testing research yielded a couple different frameworks that are available for use. The viable ones that I found for us to choose from are react-testing-library and Enzyme. This is something I did not not complete and still needs more research into. Finally, typescript analysis led me to believe that it will be worthwhile on this project. Doing this will allow for type safety on our variables and will allow us to use OOP things like classes. The downside to using typescript is that I sometimes find it to type certain things so it might have a bit of a learning curve. However, that can be negated as typescript is a superscript of javascript and can be substituted when necessary to get things to wiek.

Andrew Smith

This week I worked on researching into the Amazon RDS and EC2 technologies.

EC2: is an AWS service that provides a virtual instance of a server for our use. With these servers we can configure most settings such as: security, networking, and managing storage. It will dynamically scale up and down with the use of traffic so we don't have to worry about traffic size. Each instance of an EC2 server come with an operating system and additional software preinstalled. With the free tier we get 750 hours a month for 12 months. 750 hours comes just above 31 days a month so for our purposes we shouldn't need to worry about any fees.

Amazon RDS: is a web service that makes it easier to set up, operate, and scale a relational database in the AWS Cloud. It will manage backups, software patching, automatic failure detection, and recovery. It works with both MySQL and PostgreSQL databases so if we need to fall back on just MySQL it can. It allows for all the same features as if we were hosting the database on our own machines. With the free tier we get 750 hours a month for 12 months. The downside is that we only get 20GB free for storage and 20GB for database backup images. We don't know how much data we are getting so don't know if there will be any charges for this service. We should be able to host this on our EC2 server.

Omair Ijaz

This week was focused on creating the local database and importing the data sent from our client. The backend team focused on setting up the Java Spring Backend and creating sample endpoints that can be hosted and hit locally.

Individually, Java Spring and PostgreSQL were installed and the data that our client provided was imported into the local postgresql database. Sam created a page in our wiki that detailed how to set up the local database, so I followed those instructions. The next step was to get our database schema imported into Java Spring as entities so that we can hit the endpoints. I researched how to import databases into Spring and eventually ran into Telosys, a code generator for Java. After installing Telosys and following the instructions, I became stuck and could not make progress. After recouping with the backend team, we needed to figure out a way to generate all entities from our database. We eventually found a solution with g9, a similar piece of software.

Sam Stifter

I focused on getting access to the data for this week. We were given a database import file from the client. I set up a detailed instruction manual for the rest of the team to use so they could import it to a local instance of postgres on their development machine. This would allow us to do local development without affecting a remote database or crafty's production database. I also helped the development of the backend framework. There was a road block generating the java classes for the database tables from the PSQL database. I found a plugin for eclipse called g9 that would allow us to generate the needed files in a much simpler way than what was being attempted.

Name	Individual Contributions	Hours this Week	Hours Cumulative
Lindsey Sleeth	Research for product	9	41

	forecasting, research for React, and python vs R		
Jameel Kelley	Pipeline research, frontend testing research, typescript analysis	9	49.5
Sam Stifter	Getting access to the database with a local copy	12	53
Andrew Smith	Research into PostgreSQL and AWS hosting	7	43
Omair Ijaz	Backend code generation	8	49
Elijah Buscho	machine learning research	10	50

Plans for the Upcoming Week

(Please describe duties for the upcoming week for each member. What is(are) the task(s)?, Who will contribute to it? Be as concise as possible.)

Lindsey Sleeth
In the upcoming week I will work with Jameel and Elijah to develop our frontend solution. This includes picking up React, as well as, discussing the needs for the frontend solution with our client. We will be connecting the frontend framework to the backend to display data on the page based on queries from our client's database and creating mock data as placeholders for the results from the forecasting algorithm.

Elijah Buscho
More research needs to be done into the forecasting algorithms Lindsey has found to determine how to apply them to our situation, and which algorithms are worth using. We also need to research the specifics of the inputs to the system. Next week I plan to work on answering these questions, and also learning about React and working on the frontend.

Jameel Kelley

In the upcoming week I will be contributing to setting up CI/CD, beginning work on the frontend solution for the project, and working on mocking up data to helo ameralate communication between frontend and backend. Finally, completing frontend testing research will be on my docket.

Andrew Smith

Starting an instance of the database get it up on the server space and getting an instance of the server running on AWS.

Omair Ijaz

Individually - Research time-forecasting algorithms, research technologies for time-forecasting.
Entire Team - Set up CI/CD, schedule the rest of the semester.

Sam Stifter

Upcoming research for the technologies and forecasting. CI/CD and starter queries for the frontend.