Machine Learning / AI / Big Data Analytics

PROJECT PLAN V2

Team 32
Client: Danfoss
Adviser: Namrata Vaswani
Derek Bruun - Embedded Systems Lead / Hardware Integrations
Micky Lindsay - Software Integrations Lead / Dev Tools Manager
Smriti Manral - Lead Test Engineer / Test Designer
Victoria Rasavanh - Communications Lead / Webmaster
Jess Walters - Lead Architect / Tech Lead
http://sdmay18-32.sd.ece.iastate.edu/

Revised: 10/15

Contents

1 Introduction 3
1.1 Acknowledgement 3
1.2 Project statement 3
1.3 Operating Environment 4
1.4 Intended Users and intended uses 4
1.5 Assumptions and Limitations 4
1.6 Expected End Product and other Deliverables 4
1.6.1 Danfoss 4
1.6.2 Iowa State University Senior Design Team 5
2 Proposed Approach and Statement of Work 5
2.1 Functional Requirements 5
2.2 Constraint Considerations 5
2.3 Technology Considerations 6
2.3.1 Proposed Solution 6
2.4 Safety Considerations 7
2.5 PREVIOUS WORK AND LITERATURE 7
2.6 POSSIBLE RISKS AND RISK MANAGEMENT 8
2.7 PROJECT PROPOSED MILESTONES AND EVALUATION CRITERIA 8
2.8 PROJECT TRACKING PROCEDURES 8
2.9 OBJECTIVE OF THE TASK 8
2.10 TASK APPROACH 8
2.11 Expected Results and Validation 8
2.11.1 Planned Tests 9
3 Estimated Resources and Project Timeline 9
3.1 Personnel effort requirements 9
3.2 Other Resource Requirements 9

3.2 Financial requirements 9

3.3 Project Timeline 9

3.3.1 First Semester 9

3.3.2 Second Semester 9

4 Closure Materials 10

4.1 Conclusion 10

4.2 References 10

4.2 appendices 11

1 Introduction

1.1 Acknowledgement

The team is accepting contributions through a variety of Danfoss resources. These include the system hardware, equipment, and internal technical advice. Danfoss is providing the funding for a computational server to house the analytics engine. The company has also provided the team with VPN ready laptops to connect with Danfoss network. Technical advice is provided primarily through contacts and employees throughout the company.

1.2 Project statement

Danfoss would like to do some analytics over data they've been collecting about their assembly lines. In this project we will build a system that uses that data to predict equipment malfunction and recommend machine maintenance, using machine learning tools. From the data provided to us, we will analyze the fault pattern for the machine line and based on it determine when a failure could occur for that machine line as failure is determined by the severity of faults. "A fault is an abnormal state of a machine or system such as dysfunction or malfunction of a part, an assembly, or the whole system"(Olsson and Xiong, 2004). Since Danfoss already has sensors on machine lines, that monitors and collects data timely (putting a timestamp on it), therefore having that data available to us, our goal is to predict at time x (using data up to that time) whether the equipment will fail in the near future. As per the scope of this project we will only be looking at 1 machine line.

This project will be divided into two parts, where in the first part we will perform data analysis to analyze the pattern of faults, and in the second part we will build a system based on our analysis which will predict future failure and recommend maintenance accordingly. The system will learn over time to predict the best estimation of Time to failure for the machine line, based on early identification of unexpected events. This will allow the company to repair the machine before it breaks, saving cost and minimizing machine downtime.

Today, a large number of modern manufacturing companies are starting to incorporate data analytics and machine learning in their production and manufacturing process. This kind of predictive systems allows these companies to plan machine maintenance adaptively rather than on a fixed schedule. The purpose of such a system is to improve quality control, product quality and reduce costs, by forecasting equipment breakdowns and scheduling maintenance before they actually occur. Besides providing all these benefits this would also improve the accuracy of detecting such failures and optimize periodic maintenance operations that are carried out. Automating the manufacturing process which machine learning technique can help companies, significantly reduce the unplanned downtime and waste due to machinery failure .

1.3 Operating Environment

The analytics engine is to be housed internally within a Danfoss server. This machine is provided by Danfoss and will be connected within their live production database. The calculations produced through data analytics is utilized by a series front end displays throughout the Danfoss facility.

1.4 Intended Users and intended uses

The following parties are expected to be using the end product.

1. Danfoss Management will use the system to ensure product quantities and forecast equipment breakdowns.
2. Assembly Lines employees will view the dashboard to judge current productivity on their product lines..

1.5 Assumptions and Limitations

The following are known constraints.

1. Software and Hardware Provisions - Danfoss, as a corporate entity, must abide by both legal and internal standards when it comes to providing the team proper equipment. In order to bypass some of this, the team and client have agreed to create an isolated server system, with limited access to existing corporate data. In order to achieve this, the team will be receiving Danfoss corporate laptops, and ordering specific server hardware. Until all the hardware and software is provided, the team may not be able to do much.
2. Scheduling - All members of the senior design teams are full-time students, with some working part-time. Additionally, the Danfoss client(s) are full-time workers. Scheduling conflicts are expected to happen. When these conflicts occur, it is recommended that video conferencing substitute an in-person meeting. If no meetings are possible, alternative work should be prioritized.
3. Technology Stack - As a large, established company, Danfoss has their own set of technologies that are used internally. It is the responsibility of the senior design team to learn the technologies applicable to the project, and make sure the project itself interfaces with everything smoothly.

1.6 Expected End Product and other Deliverables

The following are deliverables that should be received by the specified parties.

1.6.1 Danfoss

By the end of the project, the Danfoss client(s) should have the following:

1. Visual Data Dashboard
    a. Easily displayed on monitors that hang over assembly line(s)
    b. Readable by workers on the assembly line floor
    c. Helps employees make better decisions related to productivity
2. Backend Module
    a. The module integration adheres to Danfoss company policy

b. The module securely consumes and interprets data across the Danfoss technology stack
c. The module integration efficiently uses Danfoss technology resources
3. Predictive Analysis / Machine Learning Algorithm
a. The algorithm predicts, within reasonable tolerances of accuracy, the estimated timeframe of maintenance needed assembly line machine
b. The algorithm designed should be able to parse through large sets of data
c. The algorithm designed should be scalable to growing Danfoss needs.
4. Documentation
a. Legal documentation specifying Intellectual Property, Consent, etc.
b. Technical documentation (Design Document and Project Plan) shall be able to provide the Danfoss team the ability to continue the support and development of the project
c. Communication logs (e-mail, Discord)

1.6.2 Iowa State University Senior Design Team

By the end of the project, the Iowa State Senior Design Team (Team 32) shall have the following for themselves:

1. Demo Web Dashboard
    a. As most of the project will remain with Danfoss, a demo dashboard may be integrated with the team's senior design website.
    b. The demo dashboard should be populated with dummy data, as to respect Danfoss' proprietary data
2. Documentation
    a. Documentation created for Danfoss shall be kept for reference

2 Proposed Approach and Statement of Work

2.1 Functional Requirements

1. The solution must interface with the existing Danfoss technology stack.
2. The solution must also work with the semi-isolated student system.
3. The solution must be able to parse large amounts of data.
4. The solution must be able to pull data from PLCs in real time

2.2 Constraint Considerations

Standard protocol will need to be based on what Danfoss defines. They, as the business client, are bound by legal specifications, and as such, we must abide by them. Danfoss primarily uses the Microsoft stack so most of our code will follow the MVC, C# and ASP.NET formatting. For databases Danfoss is using oracle databases and modifying those using TOAD, in order to simplify the transfer process we have also elected to use TOAD for our database monitoring and development. For our user interfaces we are using ignition, a guided UI tool that Danfoss has previous experience with so it should be simple to integrate with their existing systems. Danfoss has also expressed a desire for the system to interface with their PLCs so this will limit our technical solutions to the software they are currently comfortable and knowledgeable with.

2.3 Technology Considerations

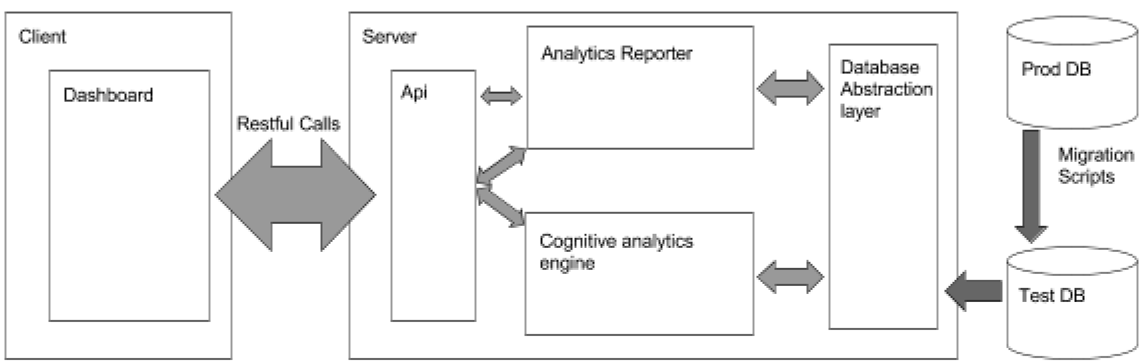
Dashboard: The proposed method for implementing the dashboard is use Ignition. Ignition is a software commonly used for making dashboards in an industrial setting. Danfoss requested that we use ignition to keep consistency across all of their dashboards. Ignition runs on Java and allows us to create our own modules to swap in and change the functionality enough to make it suit our needs.

API: The original plan was having the dashboard share an api with the machine learning and analytics engine. That plan will slightly shift with the use of ignition on the front end. Since we are now using Ignition, we will now need to have to tie the machine learning engine into the Ignition server to be able to access the dashboard to display new findings.

Analytics Reporter: The analytics reporter will be composed of an ignition module on the server. This module will access the DB. We will work this into using the abstraction layer so we can query both a mySQL test DB and a Oracle prod DB.

Cognitive Analytics Engine: Implementation using Caffe to make a robust engine to perform persistent analysis of all incoming data. Then it can output findings into the ignition server.

2.3.1 Proposed Solution



Dashboard: The dashboard is the main front end that will deliver the information to the user. This will use Ignition for the main dashboard rendering.

API: This api is the middleman between server modules and the frontend. This api can also function as an interconnect between the cognitive analytics engine and the analytics reporter.

Analytics Reporter: The analytics reporter is responsible for gathering and organizing relevant information to send to the client. This means gathering data from the DB and the analytics engine to then send to the client to display.

Cognitive Analytics Engine: This module is responsible for analyzing and interpreting the data from the assembly lines both in real-time and retrospectively. This means that we will have a constant

flow of data and be able to analyze the past and present concurrently looking for relevant patterns in the data. The data will then be accessed and interpreted by the reporter.

Database Abstraction Layer: This layer is responsible for handling and formatting our DB queries. Since the prod DB will be oracle and the test DB may not be this layer will allow us to query either DB with no extra work on the Analytics Reporter or the CAE.

Test DB: This will be a sandbox DB that we can query without the restrictions and risk of running against a prod DB.

Migration Scripts: These will be simple scripts that will pull more information from the Prod DB onto our sandbox/test DB.

Prod DB: This is the main DB containing all of the data. This DB is being constantly fed data from the production lines. We will only be using this DB to feed into a test DB until we reach a high level of maturity on our project as to avoid any possible accidents.

2.4 Safety Considerations

Our primary safety concern has to do with the security of the data passing through our systems. We will be ensuring this by using the already existing Danfoss security measures on all of our implemented solutions. Our solution will also only operate inside of Danfoss' systems to ensure that no outside networks have access to our data.

2.5 Previous Work and Literature

In past machine learning has been used in automotive plants to implement a predictive maintenance solution for a hydraulic press used in vehicle panel production. Detailed studies of the maintenance process showed that a lot of engineers' time was consumed by attending breakdowns instead of allocating resources for planned maintenance. The new solution developed using machine learning and data analytics enabled the company to predict equipment failure with an accuracy of 92% and plan maintenance more effectively. Overall equipment efficiency increased from 65% (industry average) to 85%. As a result, the optimization improved the scheduling and planning process as well as asset reliability and product quality (Romaniuk and Rutkowska, 2017).

NASA is one of the companies that has been using predictive maintenance for their turbofan engines. "Turbofan engine is a modern gas turbine engine used by the NASA space exploration agency. NASA has created the following data set to predict the failures of Turbofan engines over time. (Perera)"

Below is a diagram showing how machine learning and cloud technology is used by many manufacturing companies for predicting machine line maintenance and equipment failure.

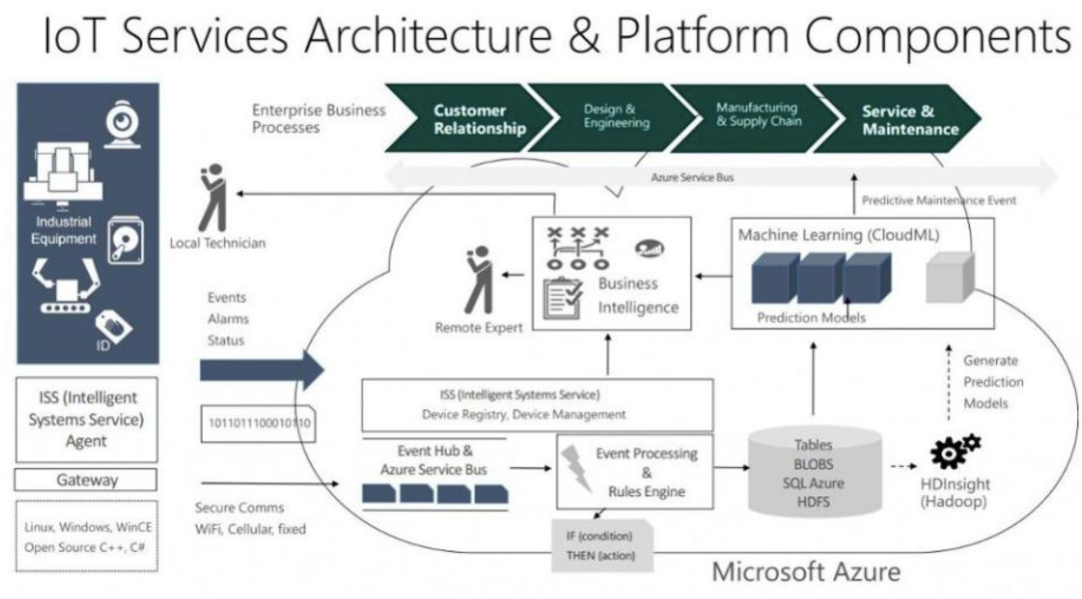


Figure 1 Ways Machine Learning Is Revolutionizing Manufacturing by Columbus, Louis. *Forbes*, Forbes Magazine, 26 June 2016

2.6 Possible Risks and risk management

1. **Team Inexperience**
2. **Unfamiliarity with Danfoss system**
3. **Assembly Line Worker Safety**

2.7 Project Proposed Milestones and evaluation criteria

1. **Project Formation & Initialization**
2. **Receive Equipment**
3. **Setup Environment**
4. **Begin Development**

2.8 Project Tracking Procedures

2.9 OBJECTIVE OF THE TASK

2.10 Task Approach

2.11 Expected Results and Validation

The validation for the project will come from comparing our old data from before the system was implemented to the data collected after the system is implemented. This will include a comparison of recalls due to faulty parts, time wasted due to broken machines and time lost due to worker error. Ideally we will do these comparisons in increments of 1 month, 6 months and yearly on a yearly basis. We will also compare the systems that have our algorithm implemented to other legacy systems that have yet to be upgraded. We will also need verification from the management

team that they are being notified of any issues on the line in a timely manner. As for the user interface we will need to both examine the data for loss of time and productivity due to human error as well as collect feedback from those who interact with the UI to ensure that it is providing them with the data in the most friendly and intuitive way possible.

2.11.1 Planned Tests

- **Perform/Add Unit tests for each new change, before pushing them remotely to github. This will ensure that the added feature and or changes works as intended and does not break our old code.**
- **Perform Integration testing. Since our application depends heavily on Danfoss's database system and their UI therefore, it is vital for us to perform integration testing as this will ensure that our system is able to communicate with other dependent systems as required.**
- **Perform functional testing to ensure that the system being designed functions as expected and satisfies the functional requirements, as a whole.**
- **Merge the branches to the master branch if and only if all the automated tests pass after pushing them to respective branches.**
- **Perform cross validation on the statistical model to evaluate and validate the model. In order to do this we will run the model over several test data set (historical data). This will be an iterative process where we will pick a subset as our test data from the data set and use the rest as training data and then repeat this process until each subset has been used as test data and record the predictions made by our model for each test data.**
- **Perform System testing.**
- **Verify that number of recalls due to fault parts has reduced.**
- **Verify workers are meeting production quotas more often.**
- **UI effectiveness - Receive feedback from workers on intuitiveness of UI**

3 Estimated Resources and Project Timeline

3.1 Personnel effort requirements

All personnel are required to:

- **Commit code**
- **participate in meetings**
- **communicate in a timely manner**

3.2 Other Resource Requirements

N/A

3.2 Financial requirements

The initial financial requirements for this project has been listed as $8000 on the original project summary.

3.3 Project Timeline

Described below is an approximate timeline for the project. According to the project summary provided to the team at the beginning of Fall 2017, the first 30% of the time should be for setting up the environment, while the remainder of the time should be for "devising our solution". The following Gantt charts are an approximation of those goals.

3.3.1 First Semester

The team will be dedicating the majority of the first semester to the following tasks:

- Familiarizing with Danfoss's existing technology stack
- Research of potential technology solutions
- Environment Setup

We have outlined our approximate schedule in this Fall 2017 gantt chart. Please contact the team if you have any trouble accessing the document.

3.3.2 Second Semester

The second semester will be devoted to development. An approximate schedule can be found here.

4 Closure Materials

4.1 Conclusion

Our project is a machine learning algorithm meant to increase the productivity of Danfoss' assembly lines. We hope to accomplish this by implementing an algorithm that will monitor both the workers and the status of the machines on the line. This will send flags to the administrators should any of the allowed tolerances not be met. Our second deliverable is a user interface to allow the workers to monitor their own progress versus the expected progress for the day. We plan on using the first semester of the course for design and rapid prototyping then moving onto the actual production version of the project in the second semester. Our final solution will be scalable to other Danfoss assembly lines.

1. Solution Goals
    a. Danfoss should have a UI that displays useful data analytics
    b. Danfoss should be able to predict when they should order machine parts, so that maintenance can be done quickly
    c. Danfoss should be able to manage their workers more efficiently, given data about what line worker is at a particular workstation at any given time.
2. Senior Design Team Goals
    a. Everyone should contribute an equal amount of work to the end product
    b. If scheduling conflicts arise, the team should be notified in advance

4.2 References

Columbus, Louis. "10 Ways Machine Learning Is Revolutionizing Manufacturing." *Forbes*, Forbes Magazine, 26 June 2016.

Romaniuk Michal, Rutkowska Barbara. "Machine Learning for Applications in Manufacturing." *Deepsense.ai Blog*, **7 Sept. 2017, blog.deepsense.ai/machine-learning-for-applications-in-manufacturing/. Accessed 23 Sept. 2017.**

Olsson, E., Funk, P. and Xiong, N. (2004). Fault Diagnosis in Industry Using Sensor Readings and Case-Based Reasoning, *Journal of Intelligent & Fuzzy Systems*, **15, (41- 46).**

Perera, Srinath. "Machine Learning Techniques for Predictive Maintenance." *InfoQ*, **www.infoq.com/articles/machine-learning-techniques-predictive-maintenance. Accessed 24 Sept. 2017.**

4.2 APPENDICES