



Master Thesis – Description of project

Course code: INMAFST-08

Credits (EC): 30 (840 hours)

To be filled in by the student and the supervisor(s) at the beginning of the final thesis research.

1. General information

Project title: Cathodic protection deterioration prediction

Student (name + signature): Jelle van Wezel

Student number: S1914146

Date: 11-12-17

First supervisor (name + signature): Michael Biehl

Second supervisor (name + signature):

External supervisor (if applicable): Richard Bosgraaf

Starting date: 11-12-17

2. Project Description:

Steel gas-pipes are placed in the ground where they form corrosion. The amount of formed corrosion is thought to be depending on multiple environment factors. For example, the composition of the soil, the ground water level, and the depth the pipes are placed. In order to prevent the formation of corrosion on the pipes a method called cathodic protection is applied. Cathodic protection uses an electrical system where anodes (metal rods) are placed in the ground near the cathode (the steel pipes). When current is introduced to the system corrosion forms on the anodes and not on the cathode.

This method is applied on multiple pipe-lines by an energy company called 'Cogas' which is located in Almelo. They have collected the environment factors by hand and know how much current is needed per pipe section. Because of the corrosion buildup over time on the anodes and the pipes, at some point in time the amount of current needed exceeds a threshold. At this point it is no longer possible or economically viable to produce the amount of current needed to prevent the corrosion. When this happens more anodes need to be placed or the steel pipes have to be replaced. Cogas wants to be able to predict when this will happen based on their collected data.

The project will be done with ValueA. ValueA provides ICT solutions for utility companies and has Cogas as a client. One of the software solutions ValueA has provided to Cogas is an interactive dashboard where pipe sections with some of their properties are shown. This project will focus on predicting when a pipe section needs maintenance by replacing the pipes or placing more anodes. These predictions could be incorporated in this dashboard.



It is unknown which environment factors (features) leads to the deterioration of the steel pipes. Methods like Supervised Variational Relevance Learning (SUVREL) and Global Metric Learning (GML) could give insight on which features are more important than others in this deterioration process. With the knowledge about the features predictions can be done on the dataset. For this project Learning Vector Quantization (LVQ) will be used to do the predictions. Its performance will be compared to the intuitive and widely used classifier K nearest neighbor (K-nn).

The cathodic protection predictions project is not the only project ValueA is involved with. Another classification problem involves Fraud detection in the energy grid. This problem forms the base of the graduation project of Sebastiaan van Loon. Since both the project of Sebastiaan van Loon and this project will be conducted at ValueA in the same period some collaboration will occur. This collaboration will consist of the development of an abstract module which includes an API that allows the existing software of ValueA to connect to a classifier implementing the abstract module.

This abstract module will be developed to ensure that the ValueA software can connect to a classification module independent of the methods the classification uses. This means that the development of the abstract module does not include the development of the actual classification algorithms and distance measures, which will be developed separately and are not part of the collaboration between Jelle van Wezel and Sebastiaan van Loon.

3. Goals:

- Research existing techniques to map LVQ and KNN on a regression problem
- Research existing methods for feature reduction/relevance
- Develop a classification module for ValueA which can predict the lifespan of steel gas pipes with LVQ, KNN and a or multiple preprocessing methods like GML and SUVREL.
- Offer a comparison of the results between the used methods
- Describe the research, used methods, developed software, and the results.

4. Working environment:

The student will be partly embedded both in the Intelligent Systems research group within the JBI, and with ValueA at the location of Cogas, a customer of ValueA located in Almelo. It is envisioned that the student will work, on average, two days a week at Cogas and the remaining days at the university and at home. The student will have a meeting at least every fortnight with the first supervisor and will have at least one stand up a week with the third supervisor and/or other employees of ValueA and will attend the monthly meetings of the Intelligent Systems research group.

The first and second supervisor will provide necessary expertise regarding the classification algorithms and dissimilarity measures used in the project. The third supervisor will provide necessary expertise about both the ValueA software solution and the energy utility market, provide data sets for the training and testing of the classifiers, and evaluate and offer feedback on the software developed.

The student will work on his own laptop and PC, and in case of failure work on a university managed system. Deployment of software will be done on a cloud environment provided by ValueA. Code revisions will be stored in a GIT repository managed by ValueA. Data which



ownership belongs to ValueA will remain in the ValueA cloud environment and will not be copied to the student's laptops unless explicit permission is granted from ValueA.

5. Deliverables:

As requested by ValueA the student will develop the software in Python.

- End of December. **Classification module description.** A document describing the classification module, its main features and its integration with the existing ValueA software.
- Start of January. **Literature survey.** A document summarizing the main findings of a literature review and the relevant references regarding Cathodic protection predictions, LVQ, KNN, Metric learning.
- End of January. **Abstract classification module.** An abstract module facilitating the connection of a classifier with the existing software and data of ValueA. (In collaboration with Sebastiaan van Loon.)
- Mid of February. **KNN classification module.** A KNN classification module for predictions of gas pipe lifespan.
- End of March. **LVQ classification module.** A LVQ module for predictions of gas pipe lifespan with relevance learning as initialization.
- Mid of April. **Test results.**
- Start of May. **Thesis draft.**
- Mid of May. **Presentation draft.**
- End of May. **Presentation to the board of Cogas.**
- End of May. **Final presentation, code, and thesis.**

6. Time planning (including end date):

Start	End	
11-12-17	27-12-17	Abstract classification module.
11-12-17	08-01-18	Literature study.
08-01-18	09-02-18	KNN classification module.
09-02-18	23-03-18	LVQ classification module.
24-03-18	09-04-18	Technical finalization, aggregating test result.
24-03-18	07-05-18	Writing of thesis, making presentation.
07-05-18	31-05-18	Final presentation, final thesis. End of project.

7. Composition of the final grade:

The final grade of the graduation project will be calculated using the following weighting.

Process: 20%

Thesis: 25%

Technical contributions: 10%

Presentation: 20%

Scientific contribution: 25%

