AI-assisted control for synthesis of solar cell materials

Degree Project E in Data Science – 1DL510

Background

One of the main aims of materials science is to discover new materials for improved technologies. Energy technologies such as batteries, solar cells and fuel cells are especially reliant on advanced materials. In developing new materials, there are many synthesis parameters to be controlled, and AI methods have huge potential to help researchers develop and optimise new materials more efficiently.

The Solar Cell Technology division at Uppsala University are currently in the process of building equipment for the investigation of a wide range of new materials intended for future solar cell and battery applications. This project aims to implement machine learning methods to assist researchers in maximising the utilisation of the equipment.

Task description

Magnetron sputtering [1] is a process used to deposit thin films of materials for a wide variety of applications, from packaging and corrosion protection to optical coatings, sensors and solar cells.

In the Solar Cell Technology division, magnetron sputtering is used to explore new interesting materials for next-generation solar cells. The first step for investigating a new material is to establish suitable process conditions for sputtering of the constituent elements. There are several goals including: a) a process which is stable over time without plasma fluctuations, b) a fast process, and c) a process which does not induce too much mechanical stress in the deposited films.

Our experimental parameters include the pressure P and flow rate F of sputtering gas (argon), as well as the power W_i applied to up to three magnetron sources. In this project, we will use machine learning in the form of Bayesian optimisation to automatically find the range of stable process conditions for a new set of 1-3 magnetron sputtering sources, by executing and analysing a series of experiments using algorithms that will be developed. The procedure should be reusable each time new sources are installed.

The project will implement Bayesian optimisation using one-by-one or batch processing of samples, where the developed algorithms will feed new sets of process parameters to the next experiment(s). In this way the process parameter space will be explored so as to delineate the regimes of "stable" and "unstable" process conditions.

We aim to address two key questions for implementing the Bayesian optimisation as described above. First, how to reduce the multivariate time series measurements to binary labels ("stable"

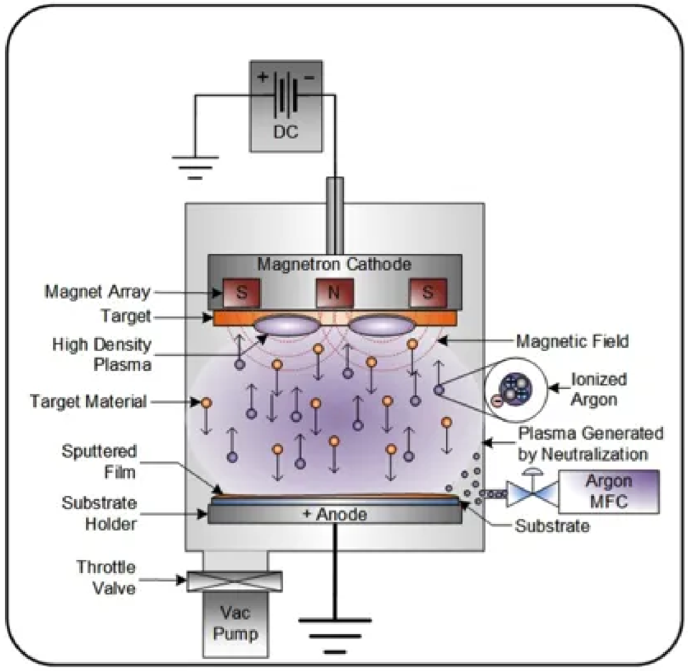

Figure 1: Schematic of a single-source magnetron sputtering system.

or “unstable” process conditions). Second, how to modify the Bayesian optimisation algorithm to identify the boundary between the two regimes rather than to optimise a performance metric.

Methods

Data for the project will be generated in the sputtering equipment of the Solar Cell Technology labs, and consists of time series measurements of voltage on each magnetron source. Methods of pre-processing the data and judging the stability of the process (based on variance in the time domain) will be developed.

Relevant courses

This project requires knowledge of machine learning, data processing, Bayesian statistics/methods and optimization, which have been acquired mainly through the courses **Statistical Machine Learning (1RT700)**, **Introduction to Data Science (1MS041)**, and **Advanced Probabilistic Machine Learning (1RT003)**. Furthermore, my mathematical foundation has been built on courses such as **Linear Algebra for Data Science (1MA330)**, **Several Variable Calculus for Data Science (1MA334)**, and **Theoretical Foundations for Data Science (1MS047)**.

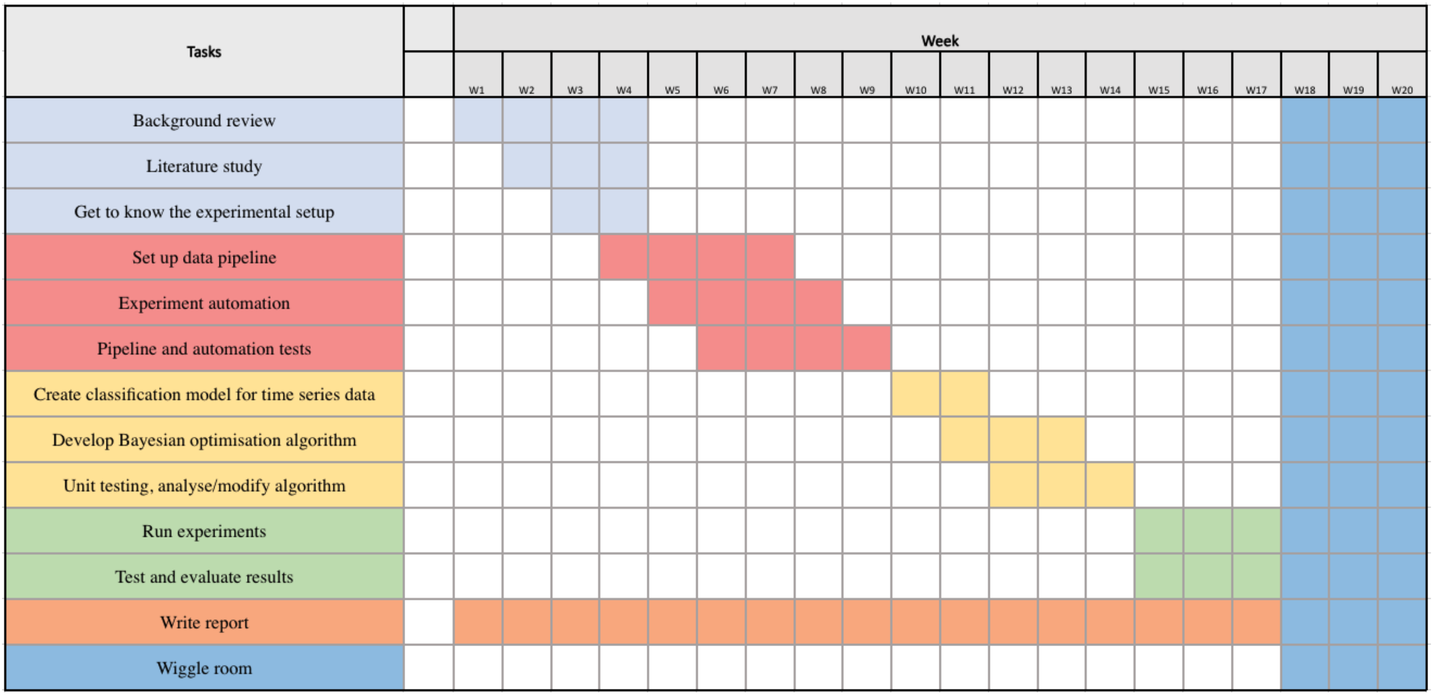

Figure 2: Weekly project plan.

Schedule

A weekly plan for this project is presented in Figure 2. We divide the project plan into color-coded blocks where we have a set of goals and milestones for each block that we aim to achieve.

Block 1: We begin the project with a background review and literature study. By the end of this block we aim to have a draft of the related works section of the project report. Additionally, time will be spent getting to know the experimental setup and document this in the methods/approach section of the report.

Block 2: In block 2 we will set up the logistics for the experiment. We will set up the data pipeline, (data collection, storage, pre-processing) and work on enabling automation such that researchers do not have to conduct experiments manually. By the end of this block we aim to have a working pipeline and to have automated the experimental setup. The work done here will also be documented in the report over the duration of the block.

Block 3: We focus on creating the classification model for the multivariate time series data, in addition to developing the Bayesian optimisation algorithm. By the end of this block we aim to have a complete implementation of the classification model and the optimisation algorithm. We will document this in the relevant parts of the report.

Block 4: In block 4 we will perform experiments in addition to analysing, testing and evaluating the results. By the end of this block, we aim to conduct some good data analysis and document it in the report.

Writing: Report writing will be done throughout the entire project in an incremental fashion instead of saving it all for the end. This streamlines the process of writing the thesis and makes progression much easier. By week 18 we aim to have a rough draft of the thesis ready.

Wiggle room: We leave the final three weeks of the project plan open to leave a margin of error.

References

[1] P. Kelly and R. Arnell, "Magnetron sputtering: a review of recent developments and applications," *Vacuum*, vol. 56, no. 3, pp. 159–172, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0042207X9900189X