**Section I: Defining the Problem**

**Scenario**

After graduating from University, you have joined a horizontally organized worker-collective organization, i.e you don’t have a boss and are part owner of the organization. The organization primarily consists of scientists who work on a number of things with electronics being more prominent. The scientists collect a lot of data on their experiments. As the data sets they collect are very large, the scientists have trouble processing them properly. Hence, you were taken on board the organization. Your main responsibility is to develop models to represent the data collected to help the scientists with their analysis. You don’t necessarily have to understand what the data exactly means, but you have to develop accurate models.

**Problem**

You have been assigned a problem recently. Your scientist colleague tells you that the problem relates to electronics where the scientists measured the variation in electron velocity between two metal plates changed in a low pressure environment with respect to the voltage applied to the plates and an external force. The sign of each quantity indicates its direction. You don’t completely understand the problem setup, but you understand that you have the data and you can work on it even if you exactly know what it means.

**Task**

The data is provided to you in the file ‘Ass1.csv’. You have to apply the machine learning process (you have learned this in the lecture and tutorials) to model the relationship between the features and the output variable. The features, ‘Voltage ’ and ‘External Force’, and the output variable, ‘Electron Velocity’ are already decided for you, and you only have to obtain the best relation between them.

For this problem you have to take:

= Electron Velocity

= Voltage

= External Force

**Section II: Required Files**

You are required to submit three files. The three files must exactly conform to the requirements below otherwise you will lose grade:

1. A python script which *only* contains a function which takes the data in Ass1.csv and returns a 1D numpy array containing all the parameters of your model (the vector from class). The function must be named ‘fit\_model1’ (the quotes are not part of the name). The output 1D-array should have coefficients in order of increasing input number and polynomial degree. For example, for the polynomial , the returned array should be array([-2.1 , -3, 9.6]). The code should be broken down into sections and with appropriate comments making it easy to follow. You may lose grade if your code is not easy to follow.
2. A python script (.py file) which contains all your code work and shows the process through which you developed your solution. The code should be broken down into sections and with appropriate comments making it easy to follow. You may lose grade if your code is not easy to follow.
3. A summary document of detailing how you arrived at your solution, what difficulties you faced and how did you try to tackle those difficulties. Your report should follow the template provided and the first line of your report should state your obtained model.

**You will receive a Fail grade by default if you don’t submit all three files listed above.**

**Section III: Our Expectations**

This assignment goes through the process we cover from weeks 1 -4. If you have been putting work in, we expect that this problem will interesting and a little challenging for you but not very difficult. What we are looking for is that you know how to tackle such a problem and know the process through which to approach the solution. You should develop your solution in a thoughtful manner and identify the interesting features of the problem. Files 2 and 3 give you the opportunity to do exactly this. Working solutions which do not give insight how they were arrived at will not be marked favorably. Moreover, banal/’professional’ and jargon-laden writing which does not convey much meaning or insight and mostly conveys straightforward information/talking-points will also not be judged favorably.

The reason we have strict requirement of what kind of files we want is because we all want them in the same form so we can check them through running our own code. If the code you provide is not exactly complies with our requirements we will have to write a special code just for your file to test it and we cannot do that for all students. Moreover, in the real-world you are usually given specific details on how the desired result will look like and if it doesn’t look like that, there can be huge issues. Just imagine a contractor builds a powerplant in your country, but the output voltage and current of the powerplant are not compatible with your country’s voltage and current. It would be a disaster!

You can go beyond what we have taught you (in fact we encourage it!!) but you should state why and how you did that.

**Section IV: Marking Criteria**

If you meet he minimal requirements for submission above, you can achieve the following grades

**Pass:** To achieve a pass mark you must show a basic understanding of the process of solving the given problem and have a basic working solution which gives a reasonable error.

**Credit:** To achieve a credit you must show a good understanding of the process and have a working solution with a reasonable error. Moreover, you can partially identify the interesting features of the problem, and some difficulties you faced and how to tackle them.

**Distinction:** To achieve a distinction you must show an excellent understanding of the process and have a well-implemented working solution with low error. Furthermore, you can identify the majority of the interesting features of the problem, the implementation issues the requirements pose and how you solved all these issues.

**High Distinction:** Everything in distinction but at an outstanding level and even going beyond!