POLITECHNIKA GDAŃSKA

<nazwa wydziału>

Imię i nazwisko studenta: <imię, nazwisko>
Nr albumu: <nr albumu>
Studia pierwszego stopnia
Forma studiów: <stacjonarne/niestacjonarne>
Kierunek studiów: <nazwa kierunku studiów>
Specjalność/profil: <nazwa specjalności/nazwa profilu>

Jeżeli praca dyplomowa licencjacka jest realizowana przez co najmniej 2 studentów, to informacje o współautorach pracy
Imię i nazwisko studenta: <imię, nazwisko>
Nr albumu: <nr albumu>
Studia pierwszego stopnia
Forma studiów: <stacjonarne/niestacjonarne>
Kierunek studiów: <nazwa kierunku studiów>
Specjalność/profil: <nazwa specjalności/nazwa profilu>

PRACA DYPLOMOWA LICENCJACKA

Tytuł pracy w języku polskim: <tytuł pracy>

Tytuł pracy w języku angielskim: <title of thesis>

Potwierdzenie przyjęcia pracy	
Opiekun pracy	Kierownik Katedry/Zakładu
podpis	*podpis*
<Tytuł, stopień, imię i nazwisko>	<Tytuł, stopień, imię i nazwisko>

Data oddania pracy do dziekanatu:

POLITECHNIKA GDAŃSKA
<nazwa wydziału>

<Logo Wydziału>
(max. wys. logo 15 mm)

OŚWIADCZENIE

Imię i nazwisko: <imię, nazwisko>
Data i miejsce urodzenia: <data, miejsce urodzenia>
Nr albumu: <nr albumu>
Wydział: <nazwa wydziału>
Kierunek: <nazwa kierunku>
Poziom studiów: <pierwszy/drugi>
Forma studiów: <stacjonarne/niestacjonarne>

Ja, niżej podpisany(a), wyrażam zgodę/nie wyrażam zgody* na korzystanie z mojej pracy dyplomowej zatytułowanej: <tytuł>
do celów naukowych lub dydaktycznych.[1]

Gdańsk, dnia ....
podpis studenta

Świadomy(a) odpowiedzialności karnej z tytułu naruszenia przepisów ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (Dz. U. z 2006 r., nr 90, poz. 631) i konsekwencji dyscyplinarnych określonych w ustawie Prawo o szkolnictwie wyższym (Dz. U. z 2012 r., poz. 572 z późn. zm.),[2] a także odpowiedzialności cywilno-prawnej oświadczam, że przedkładana praca dyplomowa została opracowana przeze mnie samodzielnie.

Niniejsza praca dyplomowa nie była wcześniej podstawą żadnej innej urzędowej procedury związanej z nadaniem tytułu zawodowego.

Wszystkie informacje umieszczone w ww. pracy dyplomowej, uzyskane ze źródeł pisanych i elektronicznych, zostały udokumentowane w wykazie literatury odpowiednimi odnośnikami zgodnie z art. 34 ustawy o prawie autorskim i prawach pokrewnych.

Potwierdzam zgodność niniejszej wersji pracy dyplomowej z załączoną wersją elektroniczną.

Gdańsk, dnia ....
podpis studenta

Upoważniam Politechnikę Gdańską do umieszczenia ww. pracy dyplomowej w wersji elektronicznej w otwartym, cyfrowym repozytorium instytucjonalnym Politechniki Gdańskiej oraz poddawania jej procesom weryfikacji i ochrony przed przywłaszczaniem jej autorstwa.

Gdańsk, dnia ....
podpis studenta

*) niepotrzebne skreślić

---

[1] Zarządzenie Rektora Politechniki Gdańskiej nr 34/2009 z 9 listopada 2009 r., załącznik nr 8 do instrukcji archiwalnej PG.

[2] Ustawa z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym:

Art. 214 ustęp 4. W razie podejrzenia popełnienia przez studenta czynu polegającego na przypisaniu sobie autorstwa istotnego fragmentu lub innych elementów cudzego utworu rektor niezwłocznie poleca przeprowadzenie postępowania wyjaśniającego.

Art. 214 ustęp 6. Jeżeli w wyniku postępowania wyjaśniającego zebrany materiał potwierdza popełnienie czynu, o którym mowa w ust.4, rektor wstrzymuje postępowanie o nadanie tytułu zawodowego do czasu wydania orzeczenia przez komisję dyscyplinarną oraz składa zawiadomienie o popełnieniu przestępstwa.

CONTENTS

LIST OF FIGURES

1. DEFINING THE PURPOSE OF THE THESIS

The main objective of this paper is an attempt to use the tools provided by Data Science to analyze data from measurements made for research work in mechanical engineering.

In some applications, a conventional data analysis methods for mechanical engineering are beginning to be insufficient and a limitation to obtaining the best possible conclusions from measurements. Here it would be useful to look around at what data analysis capabilities have been developed in recent years and try to implement them for mechanical engineering. The leading branch of science focusing on data analysis in recent years is actually Data Science and it seems that using the tools that have been developed there can bring a lot of benefits to processing measurements data.

Data science methods are not limited by the existing software - one can freely shape the interface and use any statistical tools, data science tools based on python and R libraries are optimized for processing large databases, use of machine learning methods is increasingly used to make predictions based on previous experiences.

1.1. Preliminary definition of Data Science

To begin with, it seems appropriate to give an overview of what Data Science is, the origin of this branch of science and what branches of science it is related to. The following diagram, presented below on figure 1.1, can be useful for this purpose; it shows the overlap of the branches of science that data science combines.

Figure 1.1: Data Science graph. [3]
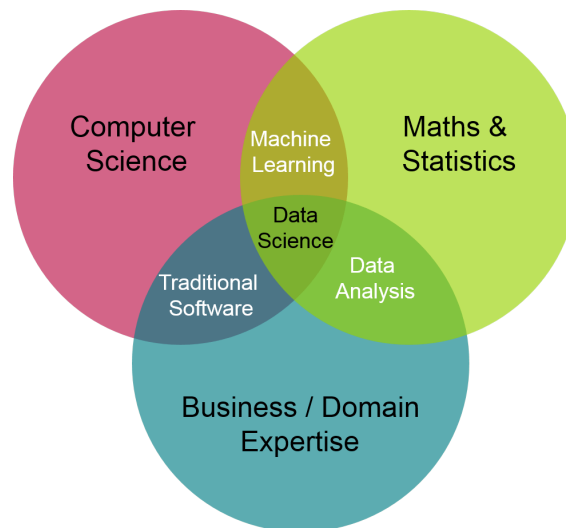
One can then recall a short definition, the unfolding of which, combined with the diagram above, should give a fairly good intuition about the complexity of this domain.

"Data science combines the scientific method, math and statistics, specialized programming, advanced analytics, AI, and even storytelling to uncover and explain the business insights buried in data." [4]

First of all Data Science is the domain of study that deals with vast volumes of data which are collected from different sources. Here we are dealing with the lowest circle in the diagram 1.1, the data can come from a business that wants to analyze its actions (for example, it can be traffic data from a store's website, or GPS data from trucks of a transport company) or from specialized research from which we want to draw the greatest insight. The data used for analysis can be from multiple sources and present in various formats.

Data Science uses modern tools and techniques that come from both circles of the top graph 1.1. Techniques derived from mathematics and statistics provide long-established capabilities for computing, analyzing, and processing data. Computer science, on the other hand, provides the ability to perform these calculations on large data sets and curate machine learning algorithms to build predictive models.

The combination of all this gives the ability to find unseen patterns, extract meaningful information and make business decisions that would not be possible with conventional solutions.

The above short definition shows that it is a field of science which takes a multidisciplinary approach and is aimed at the most accessible presentation of data in a business environment where it is often necessary to show the results of some actions to people who are not involved in a given field. Using these methods focuses on getting the most utility out of data, which in today's world is a large and often growing set.

This short introduction will be developed in one of the following chapters (Chapter 4) where aspects such as: lifecycles, tools, use cases and more will be presented in depth.

1.2. Acquisition of measurement data for analysis

Thanks to the possibility of cooperation with the department of mechanical engineering, it was possible to obtain data resulting from the work on plain bearings that they conduct. Initial discussions revealed a need to expand the range of tools for analyzing the data obtained from this research for reasons discussed in more depth in Chapter 2 of this thesis. Thanks to this possibility, it was possible to obtain real data from measurements and there was no need to create measurements or even simulate what measurements would be just to complete this paper.

Another advantage will be the possibility of comparing the results of analyses based on data science methods to analyses conducted by a unit that deals with these bearings on a daily basis, and possibly supporting their work with tools that can be selected in further stages of this work. In the case of the team that deals with these measurements, they have already been in contact with data scientists from outside the university in the form of orders to create software for processing data from specific measurements. From this point of view, it seems all the more reasonable to combine efforts to explore the topic of data analysis using data science methods.

1.3. Methodology for the analysis

The preliminary idea of the work presented in this paper is to analyze the results of measurements using selected data science methods and to compare these results with the results of analysis performed using conventional methods of mechanical measurement analysis that have been used so far by the team who perform the measurements.

2. IDENTIFYING PROBLEMS THAT ARISE WITH THE COLLECTION OF AN INCREASING AMOUNT OF MEASUREMENT DATA.

Improvements in the quality of measuring instruments and increase in their precision allow us to perform measurements with increasing accuracy and resolution. Thus, the amount of data that we obtain from measurements is also growing and it seems that this trend will continue in the coming years. Hence, there is a need to increase the possibility of processing these data, preferably in a complete free way and giving the flexibility to choose tools for their analysis from the wildest possible spectrum. In such a way as to be able to adopt to the incoming new data in case of a desire to extend measurements or change the measuring equipment and thus add new output results or possibility of adopting a different data structure.

Data processing methods, which have been used so far in the process of measurement in mechanical engineering starts to limit possibilities of extracting the greatest possible amount of informations from the measurements.

3. SEARCHING FOR AVAILABLE METHODS OF DATA ANALYSIS ALREADY USED IN MECHANICAL ENGINEERING.

When looking for a reference point, it is necessary to trace the software used so far for analyzing the measurement data. In order to narrow down the spectrum, the focus has been on the software used to date to process data from plain bearings measurements which will later be analyzed using data science methods.

3.1. Excel (Microsoft)

The first thing that comes to mind when analyzing data is, of course, Excel. The most well-known, easy-to-use and available software that is used for all kinds of analysis, processing and viewing data. The interface of the program is shown in the figure 3.1.
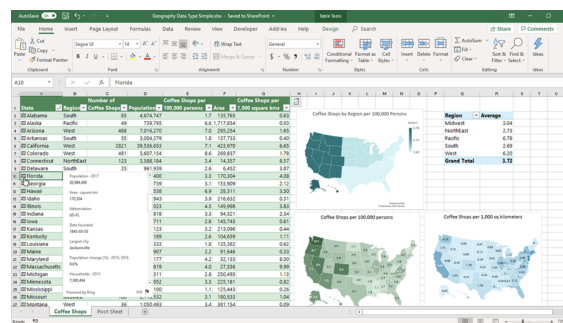

Figure 3.1: GUI of Excel software. [5]

"Microsoft Excel provides a grid interface to organize nearly any type of information. The power of Excel lies in it is flexibility to define the layout and structure of the information you want to manage. Basic tasks require no special training, and Excel allows you to work with text, numbers, and date information in a relatively open and unstructured way. Nearly 30 years after it is initial introduction, Excel remains the worlds leading spreadsheet software." **??**

3.2. Origin (OriginLab)

Origin is GUI based software aimed at scientists and therefore has a lot of capabilities often needed for measurement analysis. The software has a lot of available options, add-ons and possibilities to work with other programs. Graphic user interface of Origin software is shown on figure 3.2.

" Origin is the data analysis and graphing software of choice for over half a million scientists and engineers in commercial industries, academia, and government laboratories worldwide. Origin offers an easy-to-use interface for beginners, combined with the ability to perform advanced customization as you become more familiar with the application. Origin graphs and analysis results can automatically update on data or parameter change, allowing you to create templates
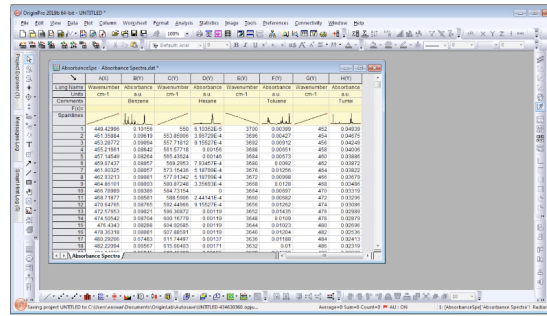
Figure 3.2: Origin GUI. [6]

for repetitive tasks or to perform batch operations from the user interface, without the need for programming. Extend the capabilities in Origin by installing free Apps available from our website." [6]

3.3. LabVIEW (NI)

LabVIEW is used to create a measurement strategy, build a test bench, and quickly view real-time data. Its capabilities are quite broad and it is a good test environment because of the ease of making changes and access to simple visualization methods immediately after measurements.
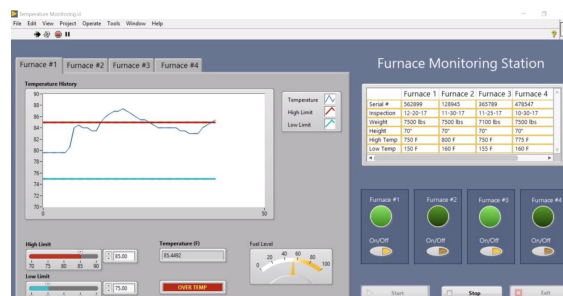

Figure 3.3: LabVIEW GUI. [7]

"LabVIEW is systems engineering software for applications that require test, measurement, and control with rapid access to hardware and data insights. LabVIEW offers a graphical programming approach that helps you visualize every aspect of your application, including hardware configuration, measurement data, and debugging. This visualization makes it simple to integrate measurement hardware from any vendor, represent complex logic on the diagram, develop data analysis algorithms, and design custom engineering user interfaces." [7]

3.4. Solartron Metrology

In this particular case, one of the software that will be used in this research work will be a detector program from Solartron Metrology. Due to the measuring devices used, the program is needed to process the output signals from the measuring devices. The signal that arrives at the computer is then processed by the software and prepared for further analysis in the form of files such as .cvs. There is a simple interface here both to control the sampling method and to adjust

10

the format of the output data. In the following discussion the role of the software from Solartron will be omitted because the focus is on the analysis of the data itself and not on the way of conducting the measurements.

4. PRESENTATION OF THE GENERAL OUTLINE OF DATA SCIENCE METHODS AND THE POSSIBILITY OF USING THEM TO ANALYZE MEASUREMENTS.

4.1. Data science

4.2. Lifecycle of data science project

When it comes to how a project is run in data science, there is a term called project lifecycle. This may differ in the way it is described as well as in the specific points depending on the source where you look for information about the project lifecycle, but in principle the main steps will be similar:

4.2.1. Understanding problem

Even before first insight into the data it is necessary to define and understand the problem which is a subject of analysis. Deep understanding of the problem principles is the most important in early stage because it can protect from loosing time for development of model which won't meet assumptions. In data science most of the projects is based on problems brought in from business or scientific demand and hence it is very important that the communication between those aggregating data and those doing strictly data science analysis is clear. Data analysis itself should not be in isolation from those doing the measuring, and there needs to be a constant exchange of insights and guidance on the desired output.

4.2.2. Data collection

Once you understand the problem, it is important to collect the right data sets. Data is the core of the whole problem and it is necessary to collect it in an appropriate way to be able to create a project based on it. In the case of this particular paper, the data is collected by researchers working in the field of plain bearings on a daily basis and on specially designed for these purposes test stands.

4.2.3. Data cleaning and processing

It is not enough to just collect data, because most of the time such raw data will not be suitable to create a model. Once collected, the data needs to be properly processed and cleaned before it can be used for the rest of the work. It is necessary to solve flaws and inconsistencies as well as deal with null values and outliers.

4.2.4.   Exploratory data analysis

4.2.5.   Model building and evaluation

4.2.6.   Communicating model results

4.2.7.   Model deployment and maintenance

4.3.   Programming language - Python

4.4.   Tools

4.5.   Use cases

5. MEASUREMENTS OF PLAIN BEARINGS AT AVAILABLE TEST STANDS.

The main measurement will focus on trying to analyze the wear of the sliding sleeve. The surface of the bushing has been asymmetrically worn down and the task of analyzing the measurement data will be to determine the volume of loss in the inner wall of the bushing.
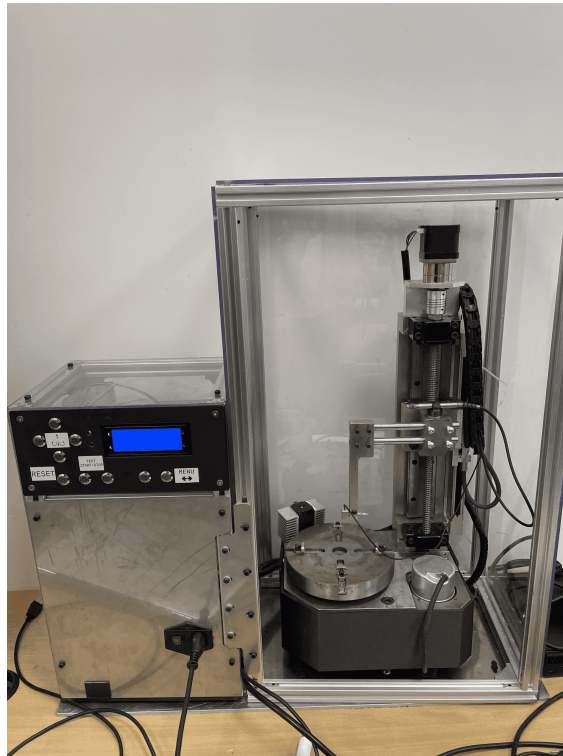

Figure 5.1: Example of measurement sleeve.

Figure 5.2: Measurement stand.

6. USE OF DATA SCIENCE METHODS TO ANALYZE OBTAINED DATA FROM PLAIN BEARING MEASUREMENTS.

7. DEVELOPMENT OF THE ANALYSIS RESULTS AND CONCLUSIONS.

BIBLIOGRAPHY

[1] A. Geron,*Hands-On Machine Learning with Scikit-Learn and TensorFlow*, O'Reilly Media, Inc., 2017

[2] P. Bruce, A. Bruce, *Practical Statistics for Data Scientists*, O'Reilly Media, Inc. 2017

[3] https://thedatascientist.com/data-science-without-programming/

[4] https://towardsdatascience.com/data-science-101-life-cycle-of-a-data-science-project-86cbc4a2f7f0

[5] https://www.microsoft.com/en-us/microsoft-365/blog/2019/02/06/excel-with-microsoft-excel-in-office-365/

[6] https://www.opengatesw.net/ms-excel-tutorials/

[7] https://www.originlab.com/index.aspx?go=PRODUCTS/Origin

[8] https://www.ni.com/pl-pl/shop/labview.html