Savitribai Phule Pune University
Board of Studies – Mechanical and Automobile Engineering
Undergraduate Program – Final Year Automation and Robotics (2019 pattern)

402046: Data Analytics Laboratory					
Teaching Scheme		**Credits**		**Examination Scheme**	
Practical	2 Hrs.	**Practical**	1	**Term Work**	50 Marks

Prerequisites: Engineering Mathematics, Artificial Intelligence & Machine Learning, Numerical and Statistical Methods, Fundamental of Mechanical Engineering

Course Objectives:
1. To explore the fundamental concepts of data analytics.
2. To understand the various search methods and visualization techniques.
3. To apply various machine learning techniques for data analysis.

Course Outcomes:

  On completion of the course, the learner will be able to

   CO1: **UNDERSTAND** the basics of data analytics using concepts of statistics and probability.

   CO2: **APPLY** various inferential statistical analysis techniques to describe data sets and withdraw useful conclusions from acquired data set.

   CO3: **EXPLORE** the data analytics techniques using various tools

   CO4: **APPLY** data science concept and methods to solve problems in real world context

   CO5: **SELECT** advanced techniques to conduct thorough and insightful analysis and interpret the results

Course Contents

Preamble:

The motivation behind the data analytics lab for mechanical engineers is to make them competent to learn data-driven decision-making involving predictive, prescriptive, descriptive, and diagnostic analytics. Data analytics offers a new paradigm of bottom-up versus top-down modelling and solving supported by the traditional physics-based approach. An engineer involved in traditional modelling (e.g., developing a finite analysis or a reliability model) looks at the problem of interest and in essence, fits in the model he/she was trained to use. An engineer equipped with data science knowledge gathers historical data and uses data-mining tools to build the model of interest. If needed, he/she can further optimize this data-driven model with tools such as evolutionary computation algorithms.

Possible approaches:

Predictive Analytics:
Predictive analytics involves the use of mathematical methods and tools such as machine learning, data mining, statistical analysis, and predictive models. It is used to:
- Identify anomalies in the process, which help in preventive maintenance.
- Estimate the demand for product, raw material etc.: based on historical data and current scenario.

scenario.
* Forecast possible outcomes based on data obtained from the process.

Prescriptive Analytics:

Prescriptive analytics is used to identify ways in which an industrial process can be improved. While predictive analytics tells when could a component/asset fails, prescriptive analytics tells what action you need to take to avoid the failure. So, you can use the results obtained from prescriptive analysis to plan the maintenance schedule, review your supplier, etc. Prescriptive analytics also helps you manage complex problems in the production process using relevant information.

Descriptive Analytics:

The core purpose of descriptive analytics is to describe the problem by diagnosing the symptoms. This analytics method also helps discover the trends and patterns based on historical data. The results of a descriptive analytics are usually shown in the form of charts and graphs. These data visualization tools make it easy for all the stakeholders, even those who are non-technical to understand the problems in the manufacturing process.

Diagnostic Analytics:

Diagnostic analytics is also referred to as root cause analysis. While descriptive analytics can tell what happened based on historical data, diagnostic analytics tells you why it happened. Data mining, data discover, correlation, and down and drill through methods are used in diagnostic analytics. Diagnostic analytics can be used to identify cause for equipment malfunction or reason for the drop in the product quality.

TERM WORK:

A] Experiments (Any 6)

Sr. No.	Data Domain	Objective	Methodology	Data type
1	Thermal / Heat Transfer / HVAC / Fluid Mechanics / Fluid Power	Predictive / Prescriptive / Descriptive / Diagnostic (but not limited to)	Statistical / mathematical /numerical/computational/intelligent (but not limited to)	Numeric or image based or data in any suitable form
2	Solid Mechanics / Design			
3	Machining / Manufacturing			
4	Automation & Robotics			
5	Maintenance / Reliability / Condition Monitoring			
6	Quality Control			
7	Materials and Metallurgy			
8	Energy Conservation and Management			
9	Industrial Engineering, Estimation, and Costing			
10	Automotive technology			

B] List of Assignments (Any Three)

The survey of methods used for data analysis in the data domain mentioned above **(Any Three)** and discussion on any case studies.

Guidelines for selection of data domain, source, size, etc.:

* The data domain must be selected from various fields of mechanical engineering such as (but

not limited to) thermal, heat power, design, manufacturing, automotive, HVAC, condition monitoring, process industry, solid and fluid mechanics, quality, materials and metallurgy, automation & robotics, energy conservation and management, ERP, Industrial engineering, estimation, and costing, etc.

- The volume of data should be considerably larger size in view of extracting meaningful insights, such as hidden patterns, unknown correlations, trends, and customer preferences through tools such as machine learning, deep learning, reinforcement learning, etc. Though the data size cannot be bluntly defined or there is no threshold, however, the data gathered from small trials/experimentation to analyse the input-output relationship should not be considered such as a trial on an external gear pump for studying its characteristics considering limited range of parameters for few trials. The appropriate data size must be selected as per the relevant data domain to yield a reliable model. For example, in the case of vibration-based condition monitoring based on numeric data, the size of data gathered depends on the sampling frequency of data acquisition and ranges from 5 kHz to 20 kHz or even more than that as per the data domain. Same for image data, the minimum number of images with appropriate resolution should be selected w.r.t data domain to yield a robust model.

- The data collected through real-time experiments is preferred however in case of no resources/facility available, data collected through simulation, survey, etc. can also be considered. The benchmark datasets made available by standard technical/academic/research/commercial/professional societies and organizations are also allowed.

- The standard instrumentation is preferred for performing experiments and data collection; however, the use of open-source hardware for building in-house low-cost data acquisition systems is also recommended.

- The choice of programming language and software depends on the data domain and the provision of the methodology used for its processing. Any standard programming language and data analytics software can be used.

- The approach mentioned above (but not limited to) should be considered while defining the problem and objectives, selecting the data domain, and deciding the methodology. The methodology can be statistical, mathematical, numerical, computational, or intelligent.

Books and Other Resources

Text Books:
1. Brunton, S. L., & Kutz, J. N. (2022). Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press.
2. Dunn, P. F., & Davis, M. P. (2017). Measurement and data analysis for engineering and science. CRC press.
3. Roy, S. S., Samui, P., Deo, R., & Ntalampiras, S. (Eds.). (2018). Big data in engineering applications (Vol. 44). Berlin/Heidelberg, Germany: Springer.
4. Middleton, J. A. (2021). Experimental Statistics and Data Analysis for Mechanical and

Aerospace Engineers. Chapman and Hall/CRC.

5. Brandt, S. (1970). Statistical and computational methods in data analysis.

6. Robinson, E. L. (2017). Data analysis for scientists and engineers. In Data Analysis for Scientists and Engineers. Princeton University Press.

7. Araghinejad, S. (2013). Data-driven modeling: using MATLAB® in water resources and environmental engineering (Vol. 67). Springer Science & Business Media.

8. Niu, G. (2017). Data-driven technology for engineering systems health management. Beijing, China: Springer.

References Books:

1. Zsolt Nagy, "Artificial Intelligence and Machine Learning Fundamentals", Packt Publishing, 2018, ISBN: 978-1-78980-165-1

2. Hastie, Trevor, Robert Tibshirani, Jerome H. Friedman, and Jerome H. Friedman. The elements of statistical learning: data mining, inference, and prediction. Vol. 2. New York: springer, 2009.

3. Zaki, Mohammed J., Wagner Meira Jr, and Wagner Meira. Data mining and analysis: fundamental concepts and algorithms. Cambridge University Press, 2014.

4. Kumar, Zindani, Davim, Artificial Intelligence in Mechanical and Industrial Engineering, CRC Press, 2021.

Assessment of Term Work

The student shall complete the above mentioned activities and prepare a Term Work in the form of Journal.

Important Note:

Term Work of the Student shall be evaluated based on the completion of experiments, group assignments and case studies. Continuous evaluation by the faculty shall be done for the award of the credit associated with the course.