# IMS065 - Data Science in Product Realization

5,0 Credits, ETC 7,5

**Elective course in Master Programs**

- Production Engineering
- Product Development


**Course Examiner**

Anders Skoogh, PhD
Professor of Production Maintenance
Director of the Production Engineering master's program
Department of Industrial and Materials Science
Division of Production Systems

E-mail: anders.skoogh@chalmers.se
Phone: +46 (0)31 772 48 06

**Course Coordinator & Administration**

Ebru Turanoglu Bekar, PhD
Researcher
Department of Industrial and Materials Science
Division of Production Systems

E-mail: ebrut@chalmers.se
Phone: +46 (0)31 772 64 13

## Course aim

The course aims to enable data-driven and facts-based decisions in mechanical engineering, specifically in the industrial product realization process. Therefore, the course aims to provide the students with fundamental knowledge about data science (including elements of Artificial Intelligence - AI and Machine Learning - ML) and abilities to apply data science techniques for improving production systems and product development. This course provides well understanding of AI systems through the appropriate formulation of the problem and the choice/application of suitable ML algorithms in order to assess the effectiveness of such algorithms using industrial case studies from product realization life cycle steps shown as in the above figure.

## Learning objectives (LOs) of the course

On successful completion of the course, the student should be able to:

LO1. Describe the fundamentals of data science, its applications (AI/ML), data-driven modelling and big data analytics.
LO2. Apply the basics of well-known libraries of the toolboxes for data scientists.
LO3. Describe steps of the data mining process.
LO4. Describe and apply visualization techniques with respect to the data mining process.
LO5. Describe and perform data pre-processing methods to ensure multi-dimensional measure of data quality.
LO6. Explain and interpret utilization of data and applicability of AI/ML algorithms for improving production systems and product development.
LO7. Interpret and discuss state-of-the-art knowledge from scientific papers related with data science in mechanical engineering.
LO8. Implement commonly used AI/ML algorithms, analyze their performance, and discuss their application using industrial applications from product realization life cycle.
LO9. Critically analyze and argue key ethical principles and potential impacts of AI on people and society and evaluate social and human requirements of systems and scenarios.

## Course content and organization

In order to reach the identified LOs above, the course is divided four modules and each module covers the following topics:

*Module 1 - Introduction to Data Science*

- Fundamentals of data science (AI/ML)
- An overview of data-driven modelling
- Introducing toolboxes for data scientists

*Module 2 - Data Mining & Visualization*

- Introduction to the data mining process and work procedures
- Plotting for exploratory data analysis (EDA)
- An overview of data quality dimensions
- Methods for data pre-processing

*Module 3 - AI and ML*

- A general introduction to AI and ML
- Examples of ML algorithms to understand in what situations they can be used
- Examples of deep learning
- Analysis of different industrial applications from product realization life cycle using AL/ML
- The ethics of AI (will be covered by reading scientific papers and discussion in literature seminar presentation)

*Module 4 - How to drive AI in your business - Project work*

- Practicing with group work project for understanding AI/ML systems through the appropriate formulation of the selected industrial cases from product realization life cycle

*Different teaching and learning activities (TLAs) will be used during the course summarized as below:*

- *Lectures* – Basis for theoretical understanding of the concepts of data science.
- *Guest lectures* – Basis for understanding different industrial applications of data science.
- *Workshops* – To reinforce the learning via more engagement with the students related to data utilization and analytics of the selected industrial cases from product realization life cycle.
- *MATLAB seminars* – To support learning of the necessary toolboxes, which are expected to use in examination project work.
- *Self-paced hands-on exercises* – Training in an interactive tutorial in MATLAB called Machine Learning Onramp course, which introduces practical ML methods, preparation for MATLAB seminars and examination project work.
- *Project work*, which is mandatory for examination, and it aims to practice skills learned throughout the course based on a structured project methodology.
- *Presentation and discussion of scientific papers* related to applications of data science in the product realization process.

*The following table also shows a summary of above TLAs for each module.*

| Module no | Lecture | Guest lecture | Workshop | MATLAB seminar | Project work | Presentation | Hands-on exercises |
|---|---|---|---|---|---|---|---|
| 1 | √ | √ | √ | | | | |
| 2 | √ | √ | √ | √ | | | √ |
| 3 | √ | √ | | √ | √ | √ (Literature seminar) | √ |
| 4 | | | | | √ | √ (Project work) | |

## Examination including compulsory elements

In this course, project-based learning is designed to give students opportunities to explore new concepts, take on new problems, formulate the problems by experiencing in different ways, and reflect on these experiences in order to improve their performance in an iterative cycle based on a structured project methodology called Cross Industry Standard Process for Data Mining (CRISP-DM). This project work is mandatory for examination and it aims to practice skills learned throughout the course.

Grading is based on the examination project work including a technical report and recorded oral presentation. Students must be approved on all assessment tasks individually such as project work, hands-on exercises, mandatory knowledge test (online quiz) through CANVAS, and literature seminar presentation to pass the course. The grading scale are Failed, 3, 4, and 5. The result of the project work has outstanding weight for grading. The result from the mandatory knowledge test will serve as decision support in borderline cases. The literature seminar and project presentations will also used as decision support for borderline grades. Further details about examination will be presented throughout the course.

## Literature

- Course PM
- Lecture materials (power-point presentations) (available at Canvas homepage of the course)
- Scientific papers (available at Canvas homepage of the course)
- Recommended resources to support learning which are provided in the following table.

Literature for reading will be separately provided for each module throughout the course. However, an overall summary is given in the following table.

| **Book information (available as e-books from Chalmers Library)** | **Recommended chapters to read (available from CANVAS page)** |
|---|---|
| Skiena S.S. (2017). The Data Science Design Manual. Texts in Computer Science. Springer, Cham. | Chapter 1 – What is Data Science? (pages 1-25)<br><br>Chapter 6 – Visualizing Data (6.1 Exploratory data analysis, pages 155-162)<br><br>Chapter 11 – Machine Learning (pages 351-390) |
| Han, J., Kamber, M., & Pei, J. (2011). Data mining concepts and techniques third edition. The Morgan Kaufmann Series in Data Management Systems. | Chapter 3 – Data Preprocessing (pages 83-124) |

**Recommended scientific papers to read (available from Chalmers Library or CANVAS page)**

1. Cao, L. (2017). Data science: a comprehensive overview. ACM Computing Surveys (CSUR), 50(3), 1-42.

2. Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. Journal of Manufacturing Systems, 48, 157-169.

3. Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. Manufacturing letters, 18, 20-23.

4. Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research, 4(1), 23-45.

5. Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining.

**Recommended link for machine learning algorithm examples i.e., classification, regression, clustering, etc. in Python**

https://scikit-learn.org/stable/

**Some helpful resources for machine learning in MATLAB**

**Machine learning in MATLAB**
https://www.mathworks.com/help/stats/machine-learning-in-matlab.html

**Some interactive apps used during the course:**
Diagnostic feature designer app link:
https://www.mathworks.com/help/predmaint/gs/import-and-visualize-data-in-diagnostic-feature-designer.html

Classification learner app video:
https://www.mathworks.com/videos/classify-data-using-the-classification-learner-app-106171.html

**MathWorks webinar series on artificial intelligence in industrial automation and machinery:**
https://www.mathworks.com/company/events/webinars/de-artificial-intelligence-in-iam.html

## Course staff

**AS: Anders Skoogh**
*Professor of Production Maintenance*
*Industrial and Materials Science*
*Chalmers University of Technology*
*anders.skoogh@chalmers.se*

**ETB: Ebru Turanoglu Bekar**
*Researcher*
*Industrial and Materials Science*
*Chalmers University of Technology*
*ebrut@chalmers.se*

**JB: Jon Bokrantz**
*Researcher*
*Industrial and Materials Science*
*Chalmers University of Technology*
*jon.bokrantz@chalmers.se*

**GA: Gauti Asbjörnsson**
*Assistant professor*
*Industrial and Materials Science*
*Chalmers University of Technology*
*gauti@chalmers.se*

**M. Siiskonen: Maria Daniela Irene Siiskonen**
*PhD student*
*Industrial and Materials Science*
*Chalmers University of Technology*
*maria.siiskonen@chalmers.se*

**KA: Knut Åkesson**
*Professor*
*Electrical engineering*
*Chalmers University of Technology*
*knut.akesson@chalmers.se*

**IG: Ida Gremyr**
*Professor*
*Industrial and Materials Science*
*Chalmers University of Technology*
*ida.gremyr@chalmers.se*

**AK: Alexander Karlsson**
*Senior Lecturer in Computer Science*
*Schools of Informatics*
*University of Skövde*
*alexander.karlsson@his.se*

**A. Syberfeldt: Anna Syberfeldt**
*Professor of Production Engineering*
*School of Engineering Science*
*University of Skövde*
*anna.syberfeldt@his.se*

**EG: Emil Gustavsson**
*Business Area Leader, Machine learning and AI*
*Systems and Data Analysis*
*Fraunhofer-Chalmers Centre*
emil.gustavsson@fcc.chalmers.se

**RA: Rohit Agrawal**
*Senior Education Customer Success Engineer*
*Nordic and Baltic Universities*
*Mathworks*
*ragrawal@mathworks.com*

**MS: Mukund Subramaniyan**
*Senior Business Analyst/Advisor*
*Insights and Data*
*Capgemini*
mukund.subramaniyan@capgemini.com

**AM: Alexander Mafi**
*Managing Consultant*
*Insights and Data*
*Capgemini*
alexander.mafi@capgemini.com

**CVR: Carl von Rosen**
*Manager Manufacturing IT/OT Systems in SKF*
carl.von.rosen@skf.com

## Schedule of the course

| W | Time | Room | Activity | Content | Responsible |
|---|---|---|---|---|---|
| **44** | **Module 1 - Introduction to Data Science** | | | | |
| 44 | 211102 13:15-15:00 Tuesday | MC | Lecture | Course introduction & Guest lecture | AS, MS, AM |
| 44 | 211102 15:15-17:00 Tuesday | HA2 | Lecture | Data science in product realization life cycle | AS |
| 44 | 211104 13:15-15:00 Thursday | ML11 | Lecture | Introduction to data science | ETB |
| 44 | 211104 15:15-17:00 Thursday | ML11 | Lecture | Project methodology - CRISP-DM | AS |
| 44 | 211105 13:15-15:00 Friday* | | - | Information about MATLAB self-paced hands on exercises, getting know MATLAB, and machine learning with MATLAB | Own work |
| ***There will be no activity on Friday since it is half day (holiday).** We suggest you to go through l*ecture slides related to MATLAB self-paced hands on exercises* and *video material related to MATLAB*, which are available on CANVAS page. | | | | | |
| **45** | ***Module 2 - Data Mining & Visualization - Part 1*** | | | | |
| 45 | 211109 13:15-15:00 Tuesday | MB | Lecture | Problem-Solution mapping - Specification of ML solutions in manufacturing | JB |
| 45 | 211109 15:15-17:00 Tuesday | MB | Workshop | Data-driven analytics in production | JB |
| 45 | 211111 13:15-14:00 Thursday | Zoom | Guest Lecture | A view on exploratory data analysis (EDA) | AK |
| 45 | 211111 14:15-17:00 Thursday | ML11 | Lecture | Industrial case study - Exploratory data analysis in product development | M. Siiskonen |
| 45 | 211112 13:15-14:00 Friday | HA2 | Community building - Q & A in IMS065 | | AS, ETB |
| 45 | 211112 14:15-15:00 Friday | HA2 | Workshop | Data utilization in product development | GA |
| -Project work hand out | | | | | |
| **46** | ***Module 2 - Data Mining & Visualization - Part 2*** | | | | |
| 46 | 211116 13:15-15:00 Tuesday | MB | Lecture | Data quality | AS |
| 46 | 211116 15:15-17:00 Tuesday | HA2 | Lecture | Methods of data pre-processing | ETB |
| 46 | 201118 13:15-15:00 Thursday | Zoom | MATLAB Seminar 1 | A demonstration of data pre-processing & feature selection with MATLAB | RA |
| 46 | 211118 15:15-17:00 Thursday | Zoom | Self-study | First part of self-paced activity in MATLAB* | ETB, Own work |
| 46 | 211119 13:15-14:00 Friday | HA3 | Community building - Q & A in IMS065 | | AS, ETB |
| 46 | 211119 14:15-15:00 Friday | HA3 | Guest lecture | Industrial case study & expert talk#1 | A. Syberfeldt |
| *You can join to zoom meeting if you have any questions regarding with the first part of self-paced activity in MATLAB. More information will be available on CANVAS page.<br><br>**-Deadline for uploading project plan to CANVAS page is Friday November 19, 23:55.** | | | | | |
| **47** | ***Module 3 - AL and ML - Part 1*** | | | | |
| 47 | 211123 13:15-15:00 Tuesday | MB | Lecture | Introduction to AI/ML/Deep learning | EG |
| 47 | 211123 15:15-17:00 Tuesday | ML1and ML4 | Lecture | Project work discussion with groups - Q & A | ETB et al. |
| 47 | 211125 13:15-15:00 Thursday | Zoom | Literature seminar | AI in industry from different perspectives* | AS, ETB |
| 47 | 211125 15:15-17:00 Thursday | Zoom | Literature seminar | AI in industry from different perspectives* | AS, ETB |
| 47 | 211126 13:15-14:00 Friday | HA3 | Community building - Q & A in IMS065 | | AS, ETB |
| 47 | 211126 14:15-15:00 Friday | HA3 | Guest Lecture | Industrial case study & expert talk#2 | KA |
| *Literature seminar contains group presentation and discussion of the selected scientific papers. More information will be available on CANVAS page. | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **48** | ***Module 3 - AL and ML - Part 2*** | | | | |
| 48 | 211130 13:15-15:00 Tuesday | Zoom | MATLAB Seminar 2 | A demonstration of modeling & evaluation with MATLAB | RA |
| 48 | 211130 15:15-17:00 Tuesday | Zoom | Self-study & Second part of self-paced activity in MATLAB* | | ETB, Own work |
| 48 | 211202 13:15-15:00 Thursday | TBD | Project work | Supervision on project work | ETB et al. |
| 48 | 211202 15:15-17:00 Thursday | TBD | Project work | Supervision on project work | ETB et al. |
| 48 | 211203 13:15-14:00 Friday | EC | Community building - Q & A in IMS065 | | AS, ETB |
| 48 | 211203 14:15-15:00 Friday | EC | Guest lecture | Industrial case study & expert talk#3 | IG |

*You can join to zoom meeting if you have any questions regarding with the second part of self-paced activity in MATLAB.
-More information related to supervision sessions will be available on CANVAS page.

| | | | | | |
|---|---|---|---|---|---|
| **49** | ***Summary of all modules*** | | | | |
| 49 | 211207 13:15-15:00 Tuesday | Zoom | Lecture | Model deployment & Operation and Maintenance | JB, CVR |
| 49 | 211207 15:15-17:00 Tuesday | Canvas/Zoom | Online quiz | Knowledge test* | AS, ETB |
| 49 | 211209 13:15-14:00 Thursday | - | Project work | Self-study | Own work |
| 49 | 211209 14:15-17:00 Thursday | - | Project work | Self-study | Own work |
| 49 | 211210 13:15-14:00 Friday | HA3 | Community building - Q & A in IMS065 | | AS, ETB |
| 49 | 211210 14:15-15:00 Friday | HA3 | Guest lecture | Industrial case study & expert talk#4 | Guest |

***The knowledge test is mandatory to attend and pass the course!**

**-Deadline for uploading project progress report to CANVAS page is Friday December 10, 23:55.**

| | | | | | |
|---|---|---|---|---|---|
| **50** | ***Module 4 - How to drive AI in your business – Project work*** | | | | |
| 50 | 211214 13:15-15:00 Tuesday | - | Project work | Self-study | Own work |
| 50 | 211214 15:15-17:00 Tuesday | MB | Lecture | Summary/recap | AS, ETB |
| 50 | 211216 13:15-15:00 Thursday | TBD | Project work | Supervision on project work | ETB et al. |
| 50 | 211216 15:15-17:00 Thursday | TBD | Project work | Supervision on project work | ETB et al. |
| 50 | 211217 13:15-15:00 Friday | - | Project work | Self-study | Own work |

-More information related to supervision sessions will be available on CANVAS page.

| | |
|---|---|
| **51 - 2** | **Holiday / Project work - The final project report and presentation must be submitted latest on Friday January 14, 23:55.** |

Week 51 (20-26/12) and Week 2 (10-16/1) are allocated to self-studies. Hence, there are no TLAs in the schedule. These two weeks are intended to be used for finalizing the project work followed by submitting the project report together with recorded oral presentation.

**Note:** All project groups are responsible for planning and executing the project work according to their own preferences.

**Note:** Most of the TLAs will be podcasted by zoom. ***A document will be separately provided in order to help you to deal with clashing course schedules***.

Please check location of rooms from the link here https://maps.chalmers.se/#a85a8be2-4ff6-4e39-9880-c2adb2a7626f

### Zoom information for the course

Most of the teaching and learning activities will be also held on line via Zoom. Please check the course calendar and join the meeting by using your Chalmers login.

**The following Zoom links will be used:**

- IMS065 Lectures (Tuesdays and Thursdays): https://chalmers.zoom.us/j/64733936536
  - Password: IMS065
- Community building & Expert talks (Fridays): https://chalmers.zoom.us/j/64213246928
  - Password: Friday!

Here is the general information regarding with the zoom.

*How to install the zoom:* https://it.portal.chalmers.se/itportal/CDAWindows/InstallZoom

*How to login via web:*

1. Go to https://chalmers.zoom.us
2. Click on Sign In
3. You will be directed to Chalmers IDP-login page
4. Log in with your "CID"

*How to login via Zoom app:*

1. Start Zoom
2. Click on Sign In
3. Click on Sign In with SSO
4. Write "**chalmers**" so it looks like "chalmers. zoom.us".
5. Click on Continue
6. You will be directed to Chalmers IDP-login page
7. Log in with your "CID"

### Connections between course elements

The following course design matrix shows the connection between course elements bases on Constructive Alignment.

| Learning Objectives (LOs) | Teaching and learning activities | | | | | | | Assestment tasks | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Lectures | Guest lectures | Workshops | MATLAB seminars | Self-paced hands on excercis | Literature seminar | Project work | Project examination report | Presentations (Literature and project work) | Knowledge test |
| LO1. Describe the fundamentals of data science, its applications (AI/ML), data-driven modelling and big data analytics | X | X | X | | | | | | | X |
| LO2. Apply the basics of well-known libraries of the toolboxes for data scientists. | | | | X | X | | X | X | X | |
| LO3. Describe steps of the data mining process. | X | | | | | | | | | X |
| LO4. Describe and apply visualization techniques with respect to the data mining process. | X | X | | | | | X | X | X | X |
| LO5. Describe and perform data pre-processing methods to ensure multi-dimensional measure of data quality. | X | | | X | X | | X | X | X | X |
| LO6. Explain and interpret utilization of data and applicability of AI/ML algorithms for improving production systems and product development. | X | X | X | | | | X | X | X | X |
| LO7. Interpret and discuss state-of-the-art knowledge from scientific papers related with data science in mechanical engineering. | X | | | | | X | | X | X | |
| LO8. Implement commonly used AI/ML algorithms, analyze their performance, and discuss their application using industrial applications from product realization life cycle. | | | | X | X | | X | X | X | X |
| LO9. Critically analyze and argue key ethical principles and potential impacts of AI on people and society and evaluate social and human requirements of systems and scenarios. | | | | | | X | | | X | |