

APPLIED MACHINE LEARNING IC

Unleashing the Power of AI

REGRESSION

Prof. Dr. Sinem Solak

Professor for AI & Technology

Program Director of BSc in AI and Sustainable Technologies

This Challenge is brought to you by

Sinem & Georg



Prof Dr Sinem Solak

🎓 Professor for AI &
Technology

📍 London, UK



Georg Bollweg

🎓 Challenge Tutor @ ToU

📍 Munich, Germany

🔍 Applied Mathematics @LMU

Kick-off Session

Agenda

- 20' Welcome
- 30' Introduction to the IC & Challenge
- 10' Wrap-Up

Welcome!

**Who is here?
Share your name,
location, and
whether you're a
program/IC learner!**



Type your answer in the Chat!



Icebreaker: “Confessions of a Future ML Expert”

“Before we dive into the magic of machine learning, let’s be real for a second...”

Pick one and complete the sentence in the chat:



The weirdest thing I wish I could predict with machine learning is...

🧠 The ML algorithm I’d take to a deserted island is... (and why!)

⚡ One thing I hope this course helps me clean up is... (my code, my data, my sleep schedule...)

⌚ In 5 years, I’ll use ML to...

Introduction to the IC

Impact Certificate

“Applied Machine Learning”

This course empowers you to harness the transformative power of machine learning.

- Go beyond theory and dive into practical, real-world applications.
- Master essential algorithms, tools, and workflows.
- Build intelligent systems that solve meaningful problems.
- Learn by doing! through projects that connect AI to sustainability and social impact.

Learning Outcomes

By the end of IC13, you will be able to:

- Explain key ML concepts
- Compare algorithms: regression, decision trees, clustering
- Select the right algorithm for your problem
- Use Python tools: [scikit-learn](#), [NumPy](#), [pandas](#), [matplotlib](#)
- Preprocess data: cleaning, transformation, feature engineering
- Train and evaluate models with real-world datasets

Introduction to the IC

The Three Challenges

Regression

Build predictive models using linear and tree-based methods



Classification

Identify categories from data using logistic regression and decision trees



Clustering

Uncover patterns without labels using unsupervised learning

Introduction to the IC

The Three Challenges – Capstones

At the end of each challenge — **Regression**, **Classification**, and **Clustering** — you'll complete a capstone project that puts your learning into practice by solving a meaningful, real-world problem using machine learning.

Each capstone will guide you through the **end-to-end ML workflow**:

1. Define the Problem

Identify a real-world challenge related to sustainability or social impact. Frame it as a regression, classification, or clustering task depending on the challenge. Clarify your goal and who benefits from your solution.

2. Prepare the Data

Source and justify a relevant open dataset. Preprocess and explore the data to understand its structure, clean missing or inconsistent values, and reflect on any ethical implications.

3. Build and Evaluate Models

Apply suitable models using libraries like **scikit-learn**. Test different algorithms (e.g., linear models, decision trees, ensemble methods, clustering techniques), tune hyperparameters, and evaluate performance using appropriate metrics (e.g., R^2 , accuracy, silhouette score).

4. Optimize and Interpret

Improve your models through regularization, feature engineering, or outlier handling. Explain key drivers, assess trade-offs, and ensure your findings are transparent and meaningful.

5. Communicate Your Results

Deliver a professional portfolio that includes:

-  A **public GitHub repo** with clean, reproducible code
-  A **technical report** explaining your workflow, results, and decisions
-  A **stakeholder-facing summary** with clear visualizations and plain-language insights
-  A short **reflection** on ethical or societal considerations

Across all capstones, your focus is on **building responsible AI solutions** that are not only technically sound, but also actionable and aligned with real-world needs.

Introduction to the IC

Choose Your Case!

Some examples:

Regression Challenge

 *Predict continuous outcomes to improve the world*

- Forecast crop yields based on soil quality and climate data 
- Predict daily energy consumption to optimize green energy supply 
- Estimate air pollution levels to inform clean transport policies 
- Model water usage in drought-prone regions to support conservation 

Introduction to the IC

Choose Your Case!

Some examples:

■ Classification Challenge

✓ *Make responsible decisions through smart categorization*

- Classify at-risk patients for early health intervention 
- Predict which buildings are energy-inefficient for retrofitting 
- Detect illegal deforestation from satellite imagery 
- Classify households eligible for clean energy subsidies 

Introduction to the IC

Choose Your Case!

Some examples:

Clustering Challenge

 *Find patterns that reveal insights and inspire action*

- Cluster urban zones by carbon emissions for targeted policies 
- Group food-insecure regions to optimize aid distribution 
- Segment recycling behavior across neighborhoods to boost engagement 
- Identify school communities with similar infrastructure needs 



Tip:

Pick a use case aligned with **your values, community needs, or global sustainability goals**.

Your project = your impact.

Introduction to the Challenge

Regression

Q1

If you could predict one sustainability or social outcome in your community, what would it be and why?



Type your answer in the Chat!

Q2

What's one real-world decision that could be improved with a better forecast?

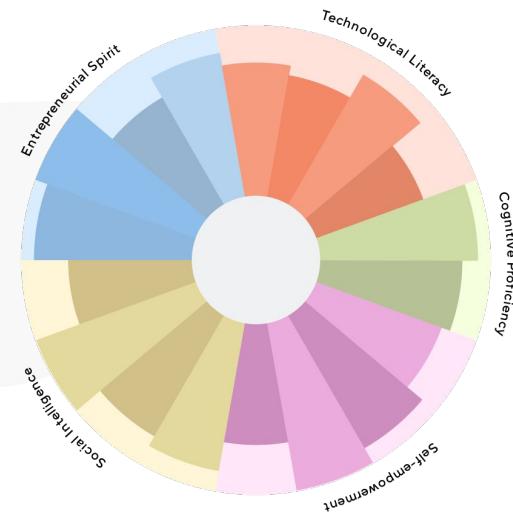


Type your answer in the Chat!

Introduction to the Challenge

Challenge Competencies

- 🎯 **Competency #1: Data Analysis** - Use data to frame a problem with clarity and purpose.
- 🎯 **Competency #2: Modelling & Simulation** - Establish the basis for applying regression techniques to a real-world use case.
- 🎯 **Competency #3: Effective use** - Prepare and structure data using industry-standard tools and workflows to enable reproducible, responsible modelling.
- 🎯 **Competency #4: Options analysis** - Compare and evaluate multiple approaches to determine the most effective model for solving a regression problem.
- 🎯 **Competency #5: Explanation & Evaluation** - Evaluate model performance, explain optimization decisions, and reflect on fairness, limitations, and impact.
- 🎯 **Competency #6: Discourse competence** - Communicate data-driven insights to diverse audiences, and reflect critically on the societal and ethical implications of your work



Introduction to the Challenge

Explore the Milestones

Milestone 1 - Kick Off Live Session 1/3

Milestone 2 - Problem Framing & Dataset Discovery

Milestone 3 - Data Understanding & Preparation

Milestone 4 - Model Development

Milestone 5 - Discussion Live Session 2/3

Milestone 6 - Model Evaluation & Optimization

Milestone 7 - Impact Reporting

Milestone 8 - Reflection Live Session 3/3

💡 Check the **assessment criteria** when preparing your capstone and engage in peer feedback before the Reflection Session.

💡 Always provide **quality feedback** on what your peer has done well and where they can improve.

Please complete necessary
milestones for the
Discussion Live Session!

Introduction to the Challenge

5' Individual Work

Go to the **Miro board** and fill out the two questions mentioned there.

5' Sharing

Share your expectations and experiences with the CO and your peers!

 Why does this challenge matter to you? How does it relate to your mission statement?

 What do you want to learn? And why?

 [LINK](#)

Wrap-Up

Wrap-Up

Organizational Stuff

Important Dates

Q&A

 Join the **Slack Channel!**
#-chl-regression

<https://tomorrowscommunity.slack.com/archives/C099QV0V8NM>

 Don't forget to **check-in for the live session** in the App!

 Get started with the **challenge**!

Wrap-Up

Key Dates of this Challenge

Discussion 

Tuesday 28.10 6.00-7.30 pm CET (90min)

Reflection 

Thursday 06.11 7-8 pm CET (60min)

Submission 

Sunday 09.11 6pm CET

Q&A