An Introduction to Artificial Intelligence (AI) in Finance

Dr. Lennart Stitz
Data Scientist
BASF Coatings GmbH

Universität Münster

Data Science Project & Presentation (1/3)

Task

- Formulate a research question based on the available data that can be addressed using supervised and/or unsupervised learning. Clearly explain the limitations of both the research question and the chosen analytical approach. Novelty is not required.
- Design and implement your solution, making all necessary decisions regarding research design (e.g., data cleaning, feature engineering, model selection). Provide well-reasoned justifications for each choice.
- Evaluate model performance and interpret the results, highlighting any limitations of the insights generated.

Data

- Annual financial data from Compustat Capital IQ to be used for the task is provided on learnweb.
- Items are defined in the provided pdf-file.

Data Science Project & Presentation (2/3)

Grading

- The final grade of the data science project is determined by the handed in report & code (67%) and the final presentation (33%).

Deadlines

- You can work alone or in groups of two. **Whether you work alone or in a group of two**, you need to **inform me via email** to lennart.stitz@basf.com **until Sunday, December 14th, 2025, 23:59.**
- All **reports** (as PDFs) **& used code** to replicated your findings needs to be handed in to lennart.stitz@basf.com by **Sunday, February 1st, 2026, 23:59.**
- All **presentations** (as PDFs or PPTs) need to be handed in to lennart.stitz@basf.com by **Wednesday, February 3rd, 2026, 23:59.**

Late or incomplete submissions will not be accepted and will result in 0 points

Data Science Project & Presentation (3/3)

Formalities

- Reports must be written in English.
- The general formatting rules of the FCM apply.
- Reports must not exceed 15 pages (including tables, figures, etc. but excluding cover page, table of contents, and references).

Guidelines & Hints

- Concise reports are preferred over lengthy narratives that lack focus.
- Select a research design that can be executed within the constraints of the available hardware. If needed, reduce the dataset size and/or narrow the optimization scope.