



# Predictive Analytics

## Organisational Issues & Use Cases



UNIVERSITY OF  
HOHENHEIM



**Intelligent Information Systems (580 A)**  
University of Hohenheim  
Institute of Interorganisational Management & Performance  
**Prof. Dr. Mareike Schoop**

# Overview of Topics

## Predictive Analytics Use Case

Decision Tree

Python Application

Support Vector Machine

Python Application

Naïve Bayes

Python Application

K-Nearest Neighbors

Python Application

Logistic Regression

Python Application

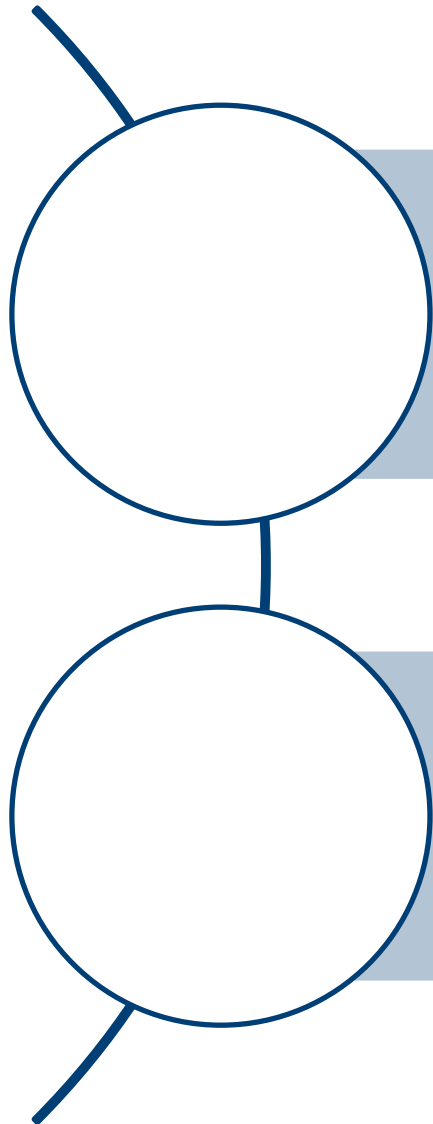
Linear Regression & Polynomial Regression

Python Application

Data Exploration, Cleansing, Processing & Transformation

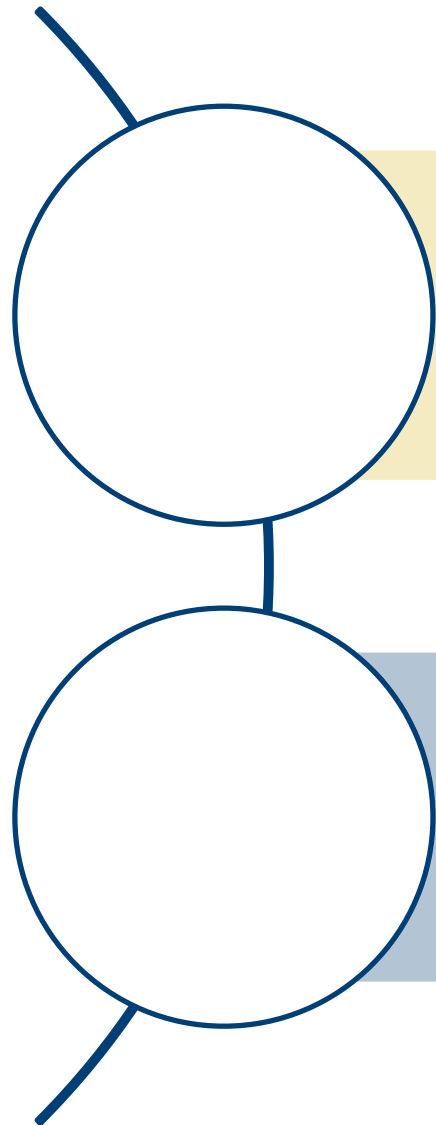
Python Application

Predictive Analytics & Project Management in Predictive Analytics



## 1. Organisational

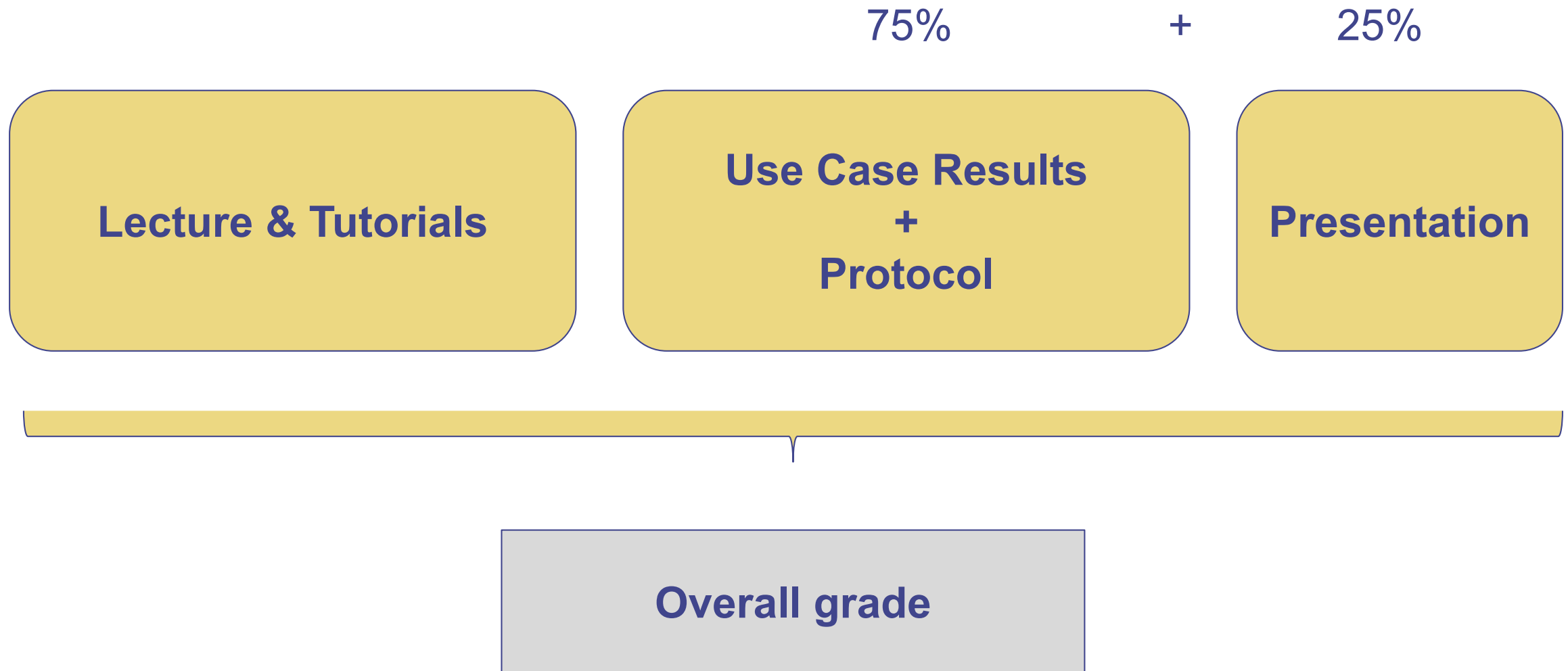
## 2. Use Cases



## 1. Organisational

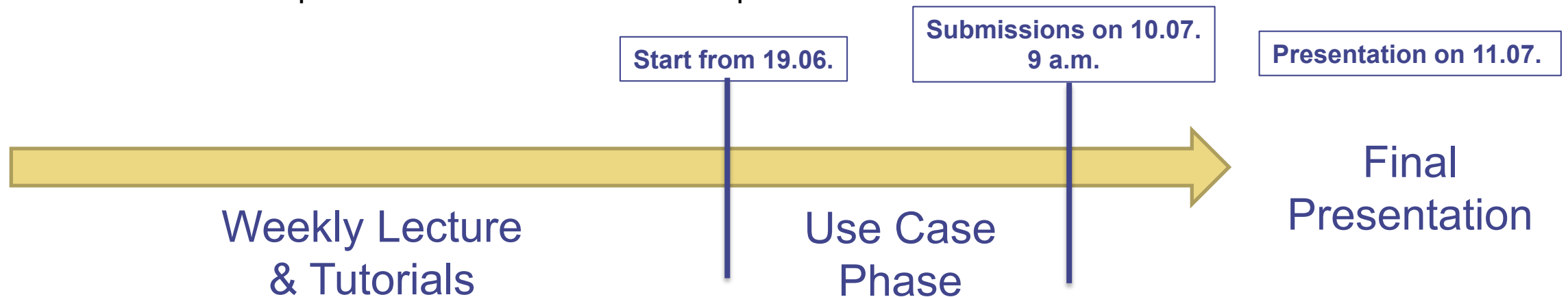
## 2. Use Cases

# Course Achievement



# Schedule & Milestones

- **Weekly lecture & Tutorial (every two weeks)**
  - Fundamental Predictive Analytics Theory
  - Practical Python implementation and show cases of selected methods
- **Practical Use Case Phase** based on the gathered knowledge (from 19.06.2023)
  - Different use cases per group
  - Application, evaluation and interpretation of ML-methods on use case data
  - Description of project implementation in form of a protocol documentation
- **Final presentation** in groups → presentation of use case results in english language
  - 20 minutes of presentation & 10 minutes of questions



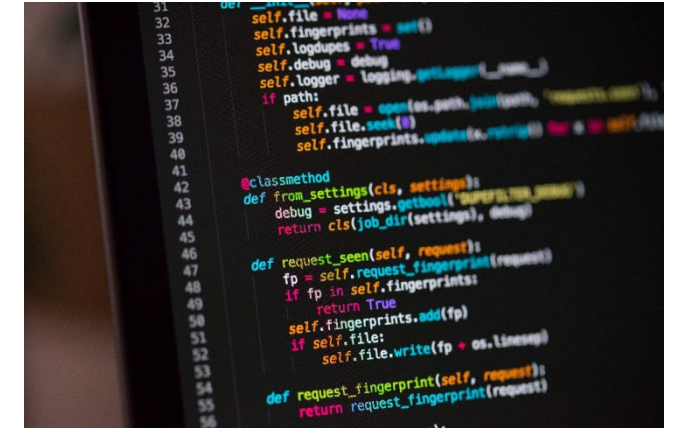
# Course Submissions I

## 1. Use Case Results as Code [Submission: 10.07.]

→ *Individual use cases which are to be solved in groups*

- Code submission
  - ◆ Data Exploration
  - ◆ Processing and Transformation of assigned dataset
  - ◆ Application and Evaluation of Machine Learning models
  - ◆ Optimisation of applied models

→ *Directly document your results in Jupyter Notebook (!)*



# Course Submissions II

## 2. Protocol of the whole Use Case phase [Submission: 10.07.]

- **Documentation** of main results
- **Interpretation** of derived results
  - ◆ From methodological point of view
  - ◆ From result-based point of view
- Methodological **reflection**
  - ◆ Challenges
  - ◆ Impressions
  - ◆ Pros and Cons

**Max. 12  
pages**



(pxhere.com)



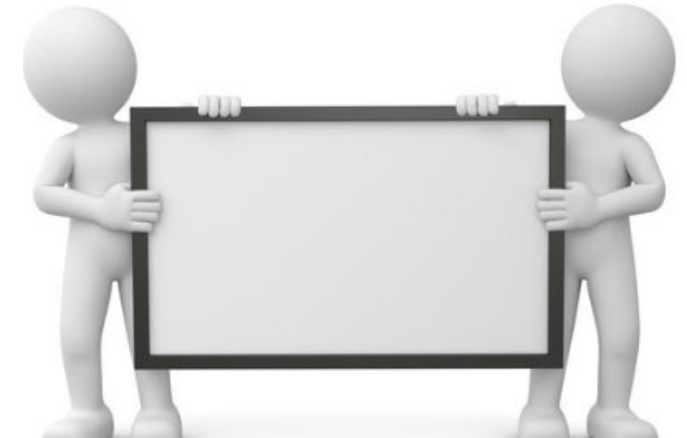
# Course Submissions III

## 3. Presentation [on 11.07.; Submission: 10.07.]

→ ***Group presentation of project results for the customer***

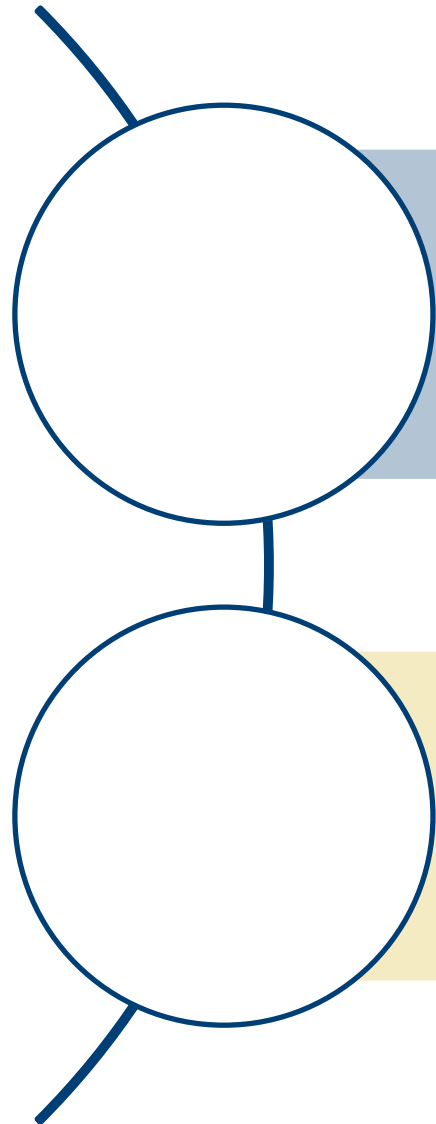
- 20 minutes of presentation and 10 minutes of question
- Content of the presentation:
  - ◆ Show how models perform
  - ◆ Present core-results of the application and evaluation
  - ◆ Give (predictive) recommendations
  - ◆ Sum up with a critical reflection
    - Impressions
    - Pros and Cons
    - Lessons Learned

Please put yourself in the position  
of a ML-engineer who works  
on the project (Use Case)



# Team Work!





## 1. Organisational

## 2. Use Cases

# General Tasks I – Data Exploration

- Please perform **data exploration** milestones to understand the case-specific data.
- Please give insights into the underlying data by using **descriptive statistics**.
- **Describe your findings** for those attributes which might play a central role in your prediction task, relativize unimportant attributes.

## General Tasks II – Data Processing

Please **clean** and **pre-process** the data if necessary, so that Machine Learning techniques can be applied in the next step.

Please apply different **data transformation techniques** to ensure processability of data.



## General Tasks III – Modeling

Please **apply** and **evaluate** different Machine Learning methods (**from Regression and Classification**) to predict the group-individual use case tasks (see specific task of your group (!)).

Please **evaluate the performance** of the models, find the **best** method-based **parameter** constellation by applying **optimisation** and conduct the **overfitting** check by applying different **validation** approaches.

Compare the different methods and **give recommendations**.

## General Tasks IV – Interpretation

- Please **interpret** your results from a **methodological** perspective and from **contextual** perspective and answer the following questions:
  - What kind of **adding-value** can Machine Learning provide with regard to the group-individual task?
  - Which Machine Learning model fits the data and the prediction task best?
  - Which Machine Learning model has a poor performance?
  - Which ML model would you recommend and why?
  - Which pattern-based information can be derived from the data?

# Use Cases



# Group 1 – Passenger Satisfaction Use Case

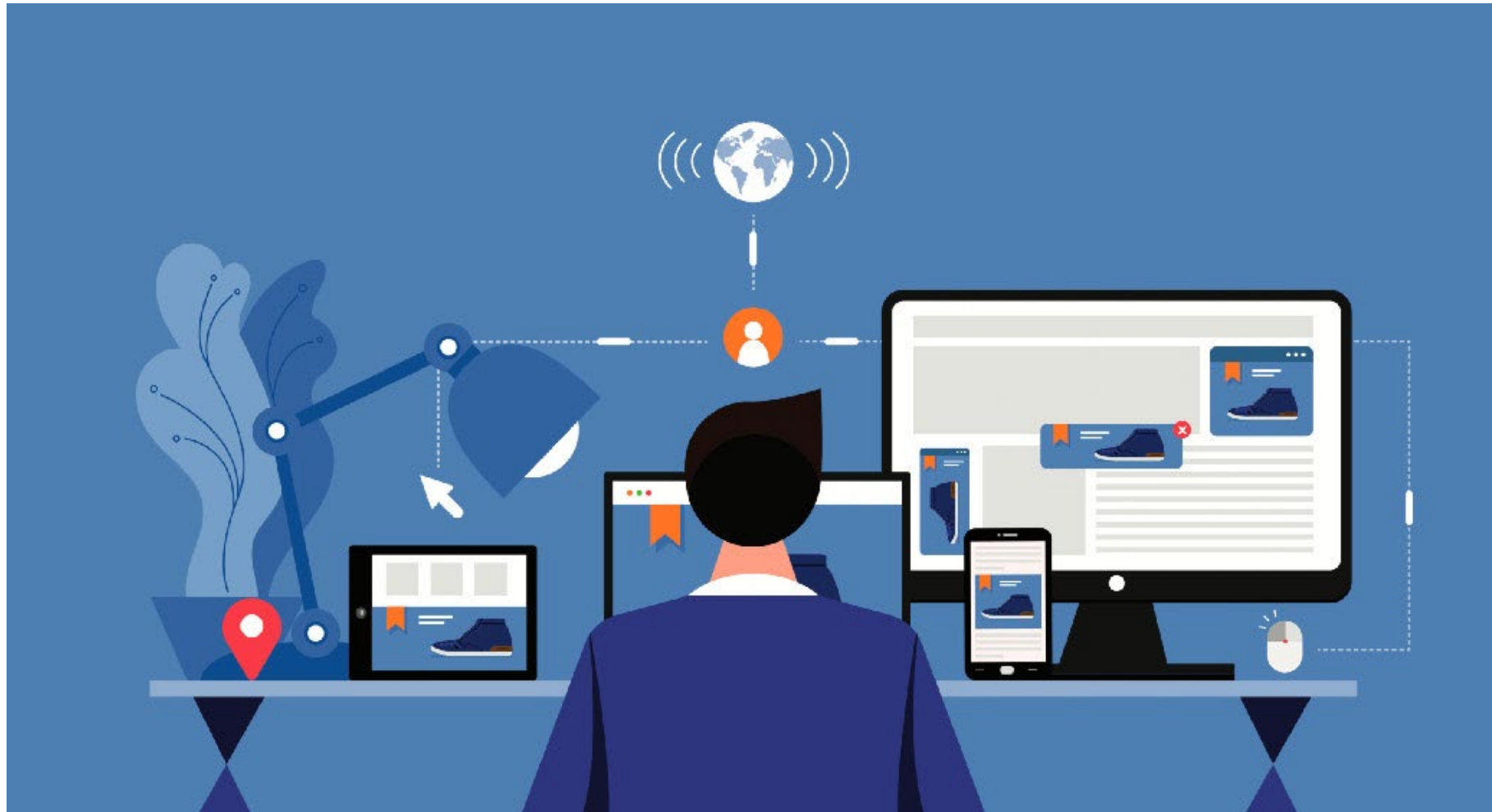


([internationalairportreview.com](http://internationalairportreview.com))

## Group 1 – Passenger Satisfaction Use Case

- *An international airways by the name **FlyAway Airlines** wants to improve the satisfaction rate of customers and reduce the delay of the flights.*
- *The management of the company wants you to derive adding-value out of the historical data they have and gives you a ML-job with following two prediction tasks:*
  1. **Please develop a ML-model which is able to predict the satisfaction level of customers.**
  2. **Please develop a ML-model which is able to predict the total delay of each flight in minutes.**

## Group 2 – Advertising Tracking Use Case



(motocms.com)

## Group 2 – Advertising Tracking Use Case

- *An international advertising company by the name **AddTec Ltd.** tracked a series of user behaviour data from a developed web-application and wants to improve the usage of functionalities.*
- *The management of the company wants to derive adding-value out of the historical data and gives you a ML-job with following two prediction tasks:*
  1. **Please build a ML-model which is able to predict whether the user will click on a specific function or not.**
  2. **Please build a ML-model which is able to predict the daily internet usage of our web-users.**

## Group 3 – Customer Churn Use Case



(clevertap.com)

## Group 3 – Customer Churn Use Case

- *An international telecommunication company by the name **TelSIM Ltd.** tracked a series of customer churn data due to the reason the retention of existing customers has become a huge challenge.*
- *The management of the company wants to derive adding-value out of the historical data and gives you a ML-job with following two prediction tasks:*
  - 1. Please build a ML-model which is able to predict whether the customers will leave the company or not.**
  - 2. Please build a ML-model which is able to predict the monthly bill (MonthlyCharge) the customer will pay.**



## Group 4 – E-Commerce Shipping Use Case



(apsfulfillment.com)

## Group 4 – E-Commerce Use Case

- *An international e-commerce company by the name **LetsBuy Ltd** wants to discover the data they had tracked with the goal of deriving strategic indicators for the company.*
- *The company has problems with the timely arrival of products and gives you a ML-job with two prediction tasks:*
  1. **Please build a ML-model which is able to predict whether the ordered product will reach on time or not.**
  2. **Please build a ML-model which is able to predict the cost of the product for the company.**



## Group 5 – Warranty Claims Use Case



(money.com)

## Group 5 – Warranty Claims Use Case

- *An international technology company by the name **TecPro Ltd.** had a series of fraud cases in warranty claims which became a huge challenge for the company due to financial issues.*
- *The management of the company wants to derive adding-value out of the historical data and gives you a ML-job with following two prediction tasks:*
  1. **Please build a ML-model which is able to predict whether the product ordering information indicates a fraud case or not.**
  2. **Please build a ML-model which is able to predict the claim value of the products.**

# Any questions?



## Next Steps

- Self-organised working in groups (!)
- **(Optional) weekly appointments for discussing open questions of your use case task**
  - **Every Tuesday between 14:00 and 15:30 in lecture hall HS 36**  
→ max. 20 minutes per group
  - Please **send me an e-mail** including the questions you want to discuss in advance and apply for the appointment if necessary until (at least) the **preceding Monday before 15:00 o'clock** → A slot will be assigned to you
  - Otherwise it will be assumed that there is no need for an appointment

## Schedule of the group-presentations (11th of July 2023 in HS 36)

Group	Use Case	Time slot
Group 3	Customer Churn	13:00-13:30
Group 2	Advertising Tracking	13:30-14:00
Group 1	Passenger Satisfaction	14:00-14:30
Group 4	E-Commerce	14:30-15:00
Group 5	Warranty Claims	15:00-15:30