



Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences



SBD3 | Find out your future salary (Group Work 2)

Prof. Dr. Branka Hadji Misheva &
Prof. Dr. Patrick Cichy

Predictive modelling project – Overview

Work on an empirical case study and derive insights from real survey data on what determines people's wages



Data Science Survey Challenge

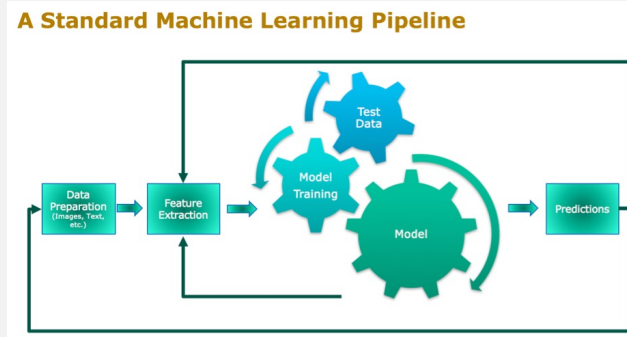
- Industry wide survey of the data science and machine learning community
- An overview of the industry on an aggregate scale
- The data contains information on people's demographics, education, experience and use of programming as well as their current salaries

Predictive modelling project– Briefing



Group Work

- Try to define very clearly what you will do and in what order
- Invest time to really understand the dataset
- Involve all (!) team members in the explorative process
- Consider what are the steps of carrying out a machine learning project



- Apply supervised learning methods that we covered in class

Predictive modelling project– Briefing

- ▶ The goal of the work is to use the data “**data_wage.RData**” and build a predictive model of **wages**.
- ▶ **Dataset overview:**
 - ▶ Subset from a real survey data
 - ▶ Dataset dimensions: 10,809 responses across 78 questions
 - ▶ Description of the individual variables is included in the file “descriptions_variables.xlsx”.

Predictive modelling project– Briefing

► Objectives of the work:

- **Build a model that will predict the variable “wage”.** Put differently, the dependent variable of the model should be $y = \text{“wage”}$.
- **You are free to decide** on the features you include as explanatory variables as well as the specific models you use. You need to provide details as to decisions you have taken.
- Once you have built your predictive model, **you need to explain it!** Identify the main features that drive the wage predictions in your model.
- **How will your model predict for you?** Use your model to predict the wage that each one from the team will earn in the future 😊
- **Coaching: 18th of April & 16th of May**

Course overview

- - Homework handout
- ★ - Group project handout

	Session														
Topic	1	2	3	4	5	6		7	8	9	10	11	12	13	14
Course overview & R recap	<div></div>														
Text Mining		<div>★</div>				<div></div>									
Advanced modelling: Overview			<div></div>												
Applied supervised learning				<div></div>	<div>★</div>										
Beyond ML: The need for explainability								<div></div>							
Coaching on group project								<div></div>			<div></div>				
Applied unsupervised learning										<div></div>					
AI governance and other current trends													<div></div>		
Students' presentations														<div></div>	

Notes:

- Easter break (no classes): 4th and 11th of April



Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences



Good luck!