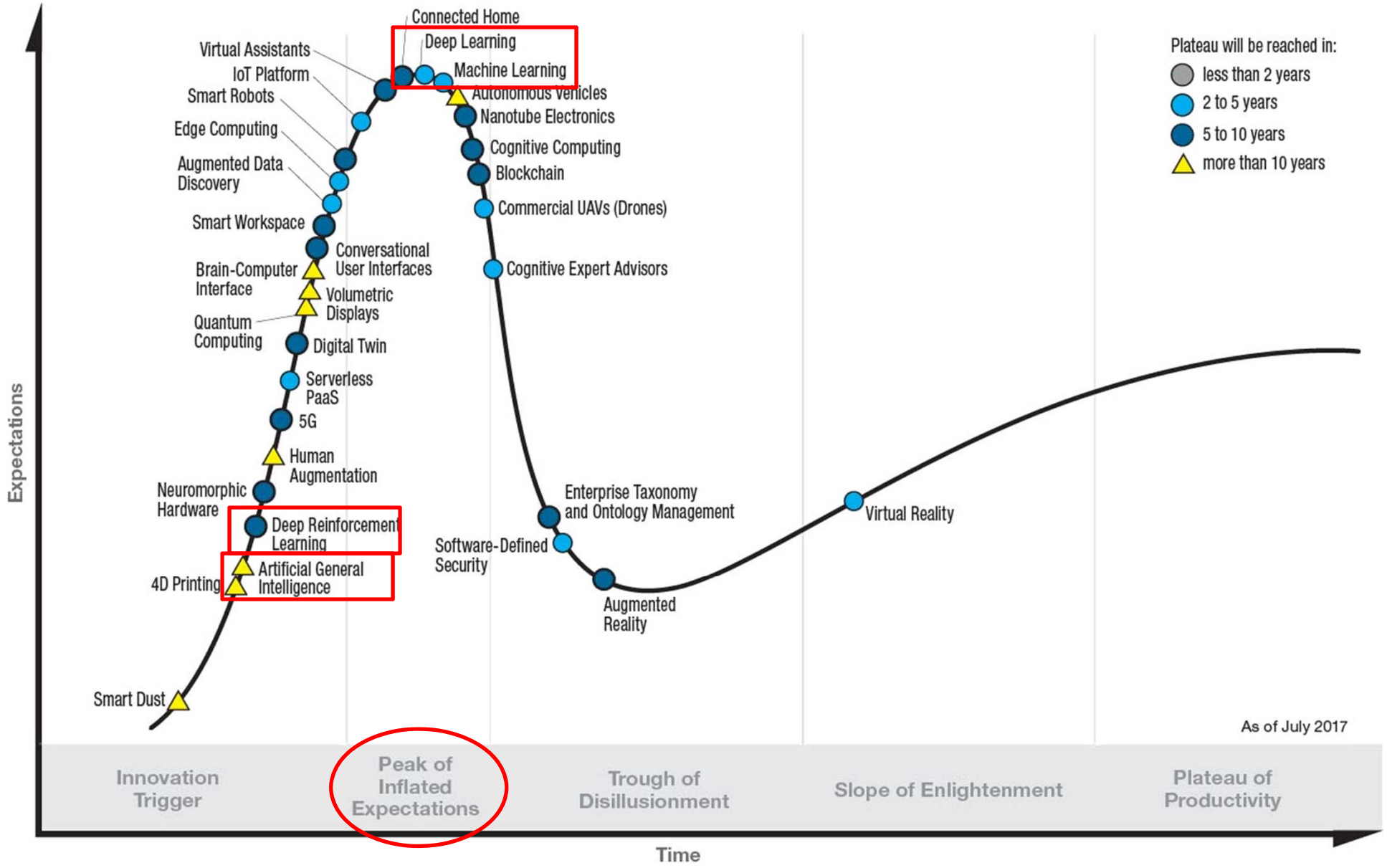
VU Machine Learning

Winter term 2019

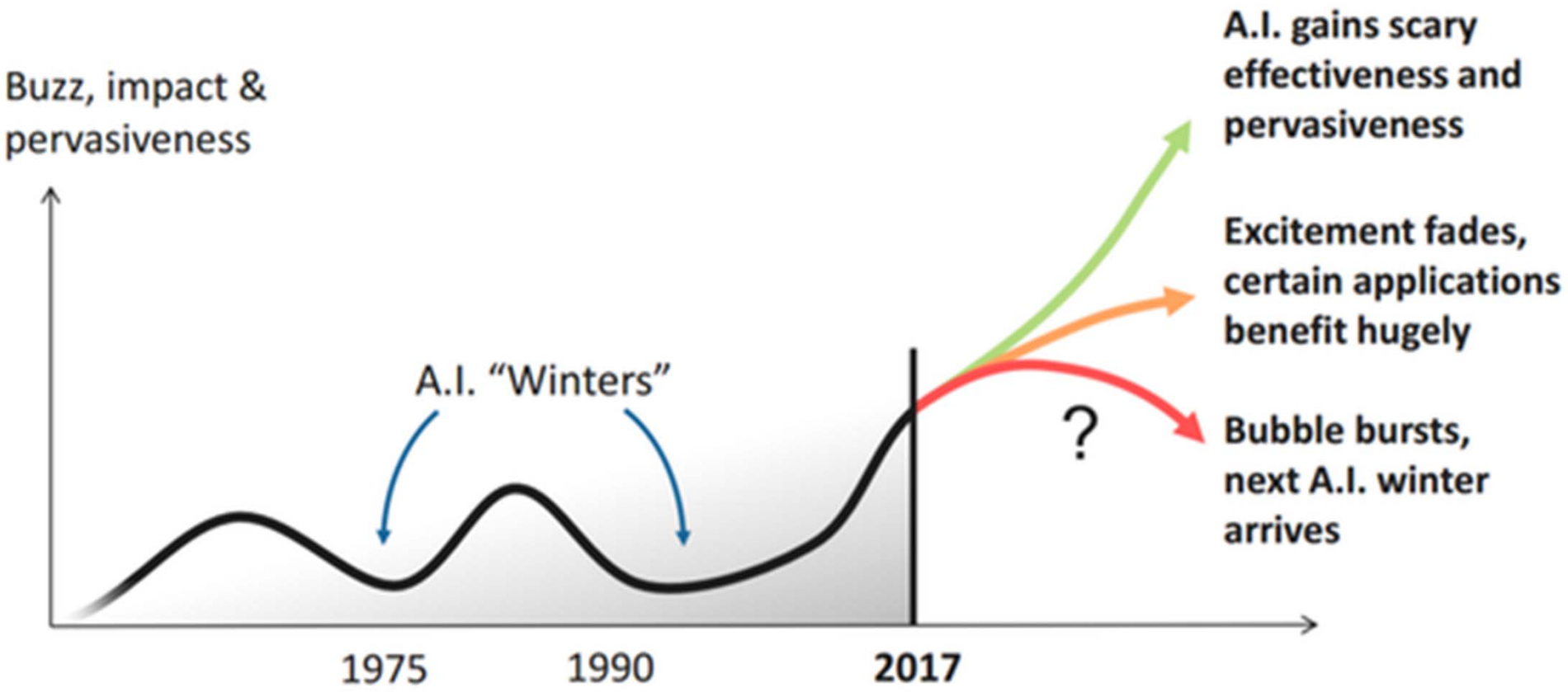
Rudolf.Mayer@tuwien.ac.at(E194), Nysret Musliu (E192)
Andreas Rauber (E194)

FAKULTÄT FÜR !NFORMATIK

Machine learning – why?

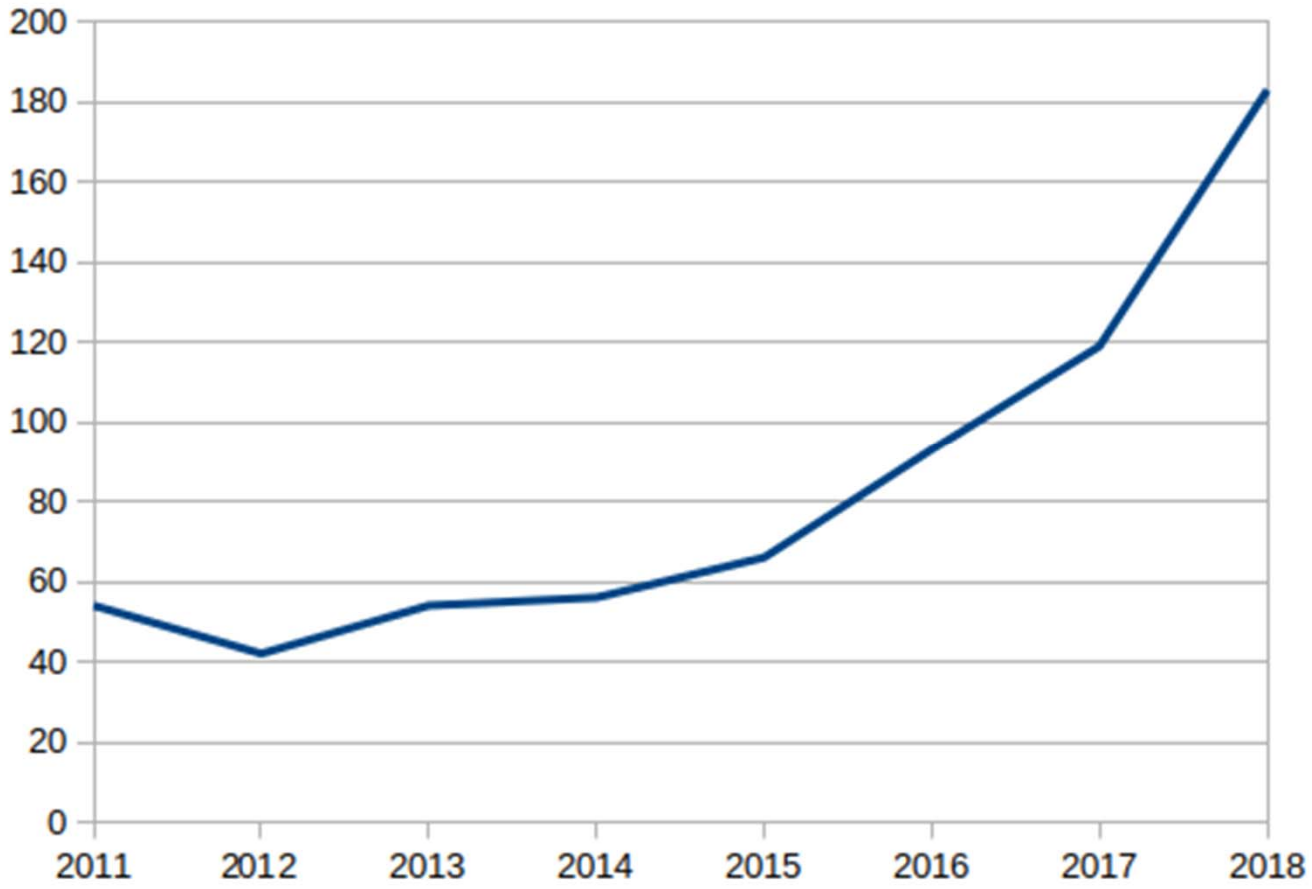


Machine learning – hype?



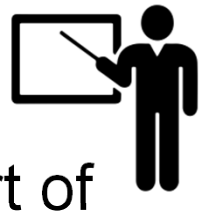
Machine learning – really popular?

- Number of students enrolled for the ML course @ TU Wien



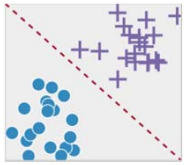
Topics

- Lecture deals with *supervised* machine learning

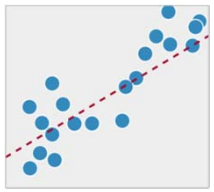


- Unsupervised learning: e.g. cluster analysis (is **not** part of this lecture)

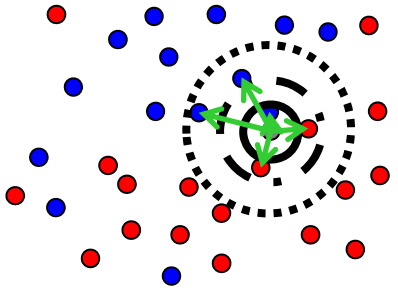
- Supervised prediction of a nominal class label: *classification*



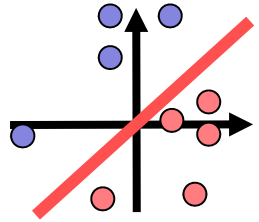
- Supervised learning to predict continuous output variables: *regression*

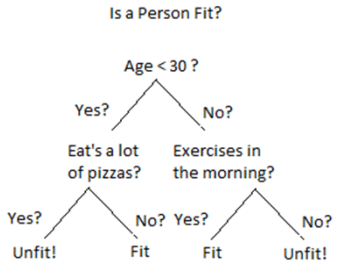


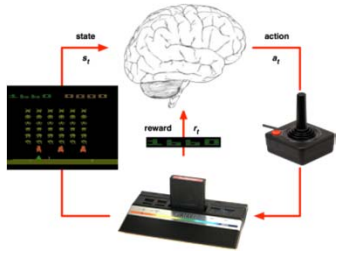
Topics



- Data types & data preparation
- Supervised machine learning algorithms
  - Decision Trees, Random Forests
  - Support Vector Machines
  - Bayesian Networks
  - Regression
  - Logistic Regression
  - Perceptron, Neural Networks, Deep Learning
- Model selection
- Automated Machine Learning
- Ensemble learning
- Significance testing







Dates

- Lectures on Wednesday, 16:00-18:00, EI 8 lecture room
  - Not every week; updated lecture schedule available in TUWEL

- 02.10. Preliminary talk
- 09.10. Introduction
- 16.10. Predicting numeric values
- 23.10. Predicting numeric values
- 30.10. k-NN, Random Forests, Decision Trees
- 06.11. Basic techniques (1R, Naïve Bayes, Covering Algorithms)
- 13.11. Support Vector Machines, Ensemble learning
- 20.11. Bayesian Networks
- 27.11. Significance testing, Feature Selection...
- 04.12. Automated ML (algorithm selection, hyperparameter optimization)
- 11.12. Deep Learning
- 18.12. Deep Learning
- 08.01. Privacy, Security, Repeatability, Explainability
- 15./22.01. Backup

Course support

- Make sure to register in TISS before the deadline [!]
  - http://tiss.tuwien.ac.at

- Register for the course in TUWEL to access material
  - https://tuwel.tuwien.ac.at/course/view.php?idnumber=184702-2019W
  - linked from TISS (automatic access when applied to lecture in TISS)

- TUWEL used for
  - Course materials (lecture slides, further readings)
  - Lab assignments
  - Forum

- TISS forum etc. will **NOT** be used !!

Target Group & Prerequisites

- Prerequisites:
  - Basic mathematics & statistics knowledge
  - Vector / matrix operations, probability density functions, etc..
  - **Programming! (one of Java, Python, R, Matlab, ...)**

- Target group:
  - Limited number of places available: 120
  - Acceptance ranked according to target group
    1. Primarily: master/PhD students in **computer science** studies
    2. Erasmus students with ML in their learning agreement
    3. Computer sciences students finishing their bachelor
    4. PhD students from other faculties, if relevant to their studies
    5. All other students
  - If you are in “category” 2-5:
    - Contact us, explain why you want to take the course

Evaluation

- Written exam
  - End of January; a total of **3 repetitions in SS 2020**
- Exercises/assignments
  - Different tasks (experiments, implementation, ...)
  - Group work of (exactly) 3 students
    - You need to form groups yourself. Start early!
    - **You are responsible for managing the work in your group**
      - Grading will be the same, even if group members leave
  - Late hand-in (max one week!): 20%-points penalty deduction
  - (i.e., you can only reach 80% of the points awarded)

- Exam and exercises contribute each 50% to overall grade
  - Each part (exam, each assignment) individually has to be passed with a positive grade !

Assignments

- Preliminary small exercise: dataset selection

- First / second assignment: Experiments and programming
  - Compare machine learning algorithms:
    - With a machine learning toolkit (e.g. R, Matlab, WEKA, Python).
    - Experiments with data pre-processing, different data sets & **classification resp. regression** techniques
    - Written report (~10 pages)
    - Discuss pre-processing, findings, differences in results per technique and datasets, patterns found, …
    - Interpret and analyse result

  - Presentations and individual discussions
  - More details after the third lecture

Assignments

- Third assignment: Experiments and programming
  - Implementation of machine learning algorithm / techniques
  - E.g., Java (WEKA), Python, Matlab, R or in other languages on request
  - Group work (can be the same as for the first assignment)
  - Presentation at the end of January
    - attendance at presentation compulsory
  - More details after the second assignment

Assignments – presentations

- You will present your findings to your peers
  - Slots of ~2-3 hours, with 8-10 groups in each slot
  - You can pick your slot in TUWEL
  - Each group member has to be present for the whole slot

- You will present two of the exercises
  - You can pick which ones
  - But there is a first-come first-serve policy
    - I.e. if all the later dates are already booked, you will **have** to present on the earlier exercises

Assignments

- Timing
  - Preliminary exercise
    - Available from Lecture 2 (9.10.), submission ~2 weeks later
  - Exercise 1:
    - Available 23.10., submission ~ 3,5 weeks later
  - Exercise 2:
    - Available 20.11. , submission ~ 3,5 weeks later
  - Exercise 3:
    - Available before Christmas

  - Presentations: in the days after each of Exercise 1/2/3

Related lectures

- Self-Organising Systems (188.413)
  - Unsupervised learning
  - Self-Organising Maps, Genetic Algorithms, Agents, …
  - Winter semester

- Information Retrieval (188.412)
  - Feature extraction from text & music, (genre) classification
  - Summer semester

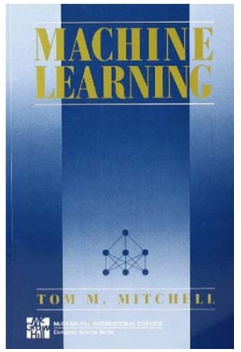
- Problem Solving and Search in Artificial Intelligence (181.190)
  - AI search techniques, learning in search
  - Summer semester

- 194.055 Security, Privacy and Explainability in Machine Learning
  - Builds on top of ML lecture
  - Summer semester (likely moved to winter semester 2019)

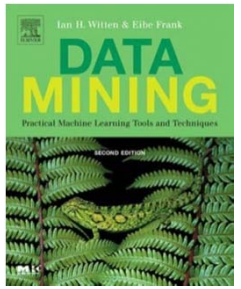
Literature

Books

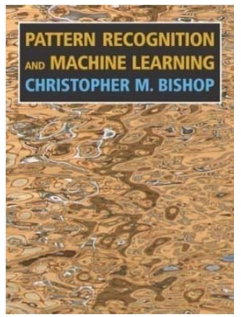
Tom Mitchell, “Machine Learning”



Ian Witten, “Data Mining: Practical Machine Learning Tools and Techniques” (WEKA Authors)



Christopher Bishop, “Pattern Recognition and Machine Learning”



TU WIEN

Questions ?

Funding opportunity

FEMTech Funding opportunity

- Funding by FFG (Austrian Research Funding Agency)
- Target group: **female students**
- Duration: 2-6 months, **paid**
- Can be used to work on a thesis or project (Praktikum)

More info: https://www.ffg.at/femtech-praktika

Interested?

- ➔ Contact Rudolf Mayer (mayer@ifs.tuwien.ac.at)
- ➔ Deadline now ☺ (October 2019)
  - ➔ Also possible later: March 2020