



university of  
 groningen

artificial  
 intelligence

# Uncertainty Estimation in Machine Learning

Introduction to the Course  
Dr. Matias Valdenegro

[m.a.valdenegro.toro@rug.nl](mailto:m.a.valdenegro.toro@rug.nl)

# Your Lecturer(s)

Dr. Matias Valdenegro  
Assistant Professor  
for Machine Learning

- Uncertainty in ML
- Computer Vision and Robotics
- Deep Learning

Open for questions,  
help, advice



# Your Lecturer(s)

Dr. Marco Zullich  
Lecturer in Machine  
Learning

- Explainable AI
- Uncertainty in ML
- Deep Learning

Open for questions,  
help, advice





# This Course

- › This course is about uncertainty estimation in machine learning models.
- › This means you will learn about models that output a probability distribution instead of a point-wise estimation.
- › The core idea is that a distribution can encode model confidence, uncertainty, and/or lack of knowledge from the model into the prediction.
- › This is useful for applications requiring safety, where it is good to know if the prediction should be trusted or not.



# Best Example - ChatGPT and LLMs

ChatGPT often produces answers that are factually incorrect, but the Chatbot believes them to be right, for example:



A father and his son are in a car accident. The father dies at the scene and the son is rushed to the hospital. At the hospital the surgeon looks at the boy and says "I can't operate on this boy, he is my son." How can this be?

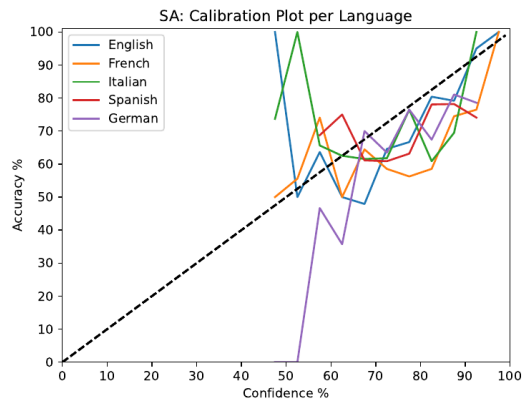


It is impossible for the surgeon to be the boy's father in this situation, as the father has already died in the car accident. The statement made by the surgeon is contradictory and therefore cannot be true.

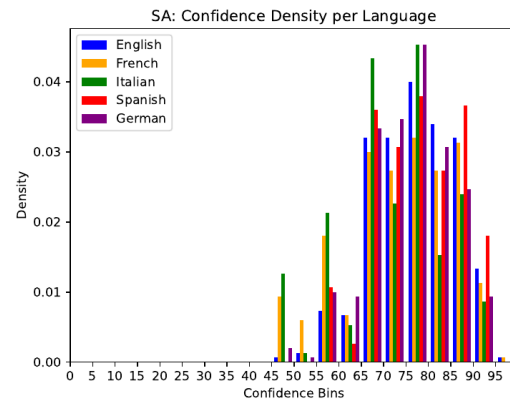
ChatGPT does not have a direct way to express confidence, if asked about confidence, usually it produces incorrect estimates of confidence, specially for wrong answers.

# Best Example - ChatGPT and LLMs

My brilliant Student Martino Pelucchi asked and evaluated ChatGPT confidence on Sentiment Analysis and Common Sense Reasoning, and the results were not good.



Calibration plot of each language for the sentiment analysis task with the diagonal line showing the perfect calibration. ChatGPT is mostly overconfident in all languages.



Density histogram of each language's confidence bins for the sentiment analysis task. ChatGPT does not give a confidence level below 50% in any language.

ChatGPT outputs no confidence less than 45%.

<https://arxiv.org/abs/2311.06427>



# Learning Objectives

- › Understand the basic concepts of Uncertainty Quantification in the context of Machine Learning.
- › Understand different architectures and methods to add uncertainty to a neural network.
- › Implement, train and evaluate Bayesian Neural Networks and models with uncertainty.
- › Perform out of distribution detection for complex tasks.
- › Evaluate quality of uncertainty through calibration and other metrics.



# Lecture & Practicals Planning

- › 25.1 – Introduction to Uncertainty in ML
- › 29.1 – Methods for UQ
- › 31.1 – Practical 1 Methods for UQ with Keras
- › 5.2 – Practical 2 Methods for UQ with Keras
- › 7.2 – Evaluation of Uncertainty
- › 9.2 – Calibration and OOD Detection
- › 12.2 – Practical on Calibration
- › 14.2 – Practical on OOD Detection
- › 16.2 – Lecture/Practical on Challenges of UQ





# Expectations

- › In our lectures, you can always ask questions, we encourage it.
- › If you do not understand something, please ask to be explained in a different way or to correct misconceptions.
- › The idea of these courses is for you to learn, so we expect you to ask questions and discuss with the lecturers.



# Evaluation and Grading

- › For a certificate of attendance, you will need to attend at least 6 instances (lectures+practicals, not including the final Q&A session/final project feedback).
- › For a certificate of competence/completion of the course, you will need to complete a small final project, whose track will be released further along the course.



# Literature

- › Unfortunately there is no book in this topic (yet).  
There are two survey papers that we can use as main reference:
  - A Survey of Uncertainty in Deep Neural Networks by Gawlikowski et al. 2021. Available at <https://arxiv.org/abs/2107.03342>
  - A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges by Abdar et al. 2020. Available at <https://arxiv.org/abs/2011.06225>
- › Each lecture links to the literature.

# Questions?

