

Data Science and Ethics

 tuyaux.winak.be/index.php/Data_Science_and_Ethics

Data science and ethics

Richting	<u>Informatica</u>
----------	--------------------

Jaar	<u>MINF</u>
------	-------------

Studiepunten	3
--------------	---

Bespreking

Dit vak is een keuzevak voor de master. Verschillende ethische aspecten voor data science worden besproken, in heel het data science proces (data gathering, preprocessing, modeling, evaluation and deployment) op verschillende vlakken (fair, transparent and accountable). Het vak is van een andere faculteit (FBE), waardoor het technisch niet heel diepgaand is, maar het is ook niet te oppervlakkig. De lessen (en het examen) vinden dan ook plaats op de stadscampus. Het bijhorend boek bevat normaal ook alle leerstof en is zeker duidelijk moest je bv. een les gemist hebben.

Het vak is niet heel moeilijk, maar vind ik wel belangrijk in onze richting om bewust te zijn van mogelijke etische gevolgen van de modellen die je maakt.

Puntenverdeling

Examen : 16/20. Presentatie: 4/20.

Het examen bestaat steeds uit twee grote vragen en vijf kleinere vragen.

Examenvragen

Academiejaar 2022-2023 (januari)

1. The company you work for wants to create a prediction model to make it more objective who to promote. The data science team gets a dataset of the employees containing their home address, gender, nationality, age, education, salary, salary when they started, years they have been working for the company and for people in the sales department also the number of sales they did. To get the accuracy as good as possible, the data from all employees working in the company is provided. (answer max 2 pages)
 1. What are possible concerns or outcries employees might have about this model? Think about the FAT principles.
 2. What are possible techniques to mitigate these concerns?
 3. What cautionary tale have we seen in the lectures related to this?

2. What are possible approaches to handle discrimination in the preprocessing phase?
Explain and demonstrate them using your own toy example containing 9 data instances with race as sensitive attribute and whether to give a credit or not as value to predict. (answer max 2 pages)
3. Small questions (max. 3 lines per question)
 1. What is the difference between 'Comprehensible' and 'explaining' (when talking about AI)?
 2. What is the difference between 'Anonimization' and 'Pseudonimisation'?
 3. What is a counterfactual explanation?
 4. What is equalized opportunity?
 5. List the six good data science practices mentioned in GDPR article 5. (just list them, you don't need to explain them).