

Final Exam for Explainable Fuzzy AI Class

Summer 2021

1. *Explainable AI.*

- 1a. What is explainable AI? Why do we need explainable AI?
- 1b. Why does it make sense to use fuzzy techniques in explainable AI?

2. *Fuzzy techniques: general description.*

- 2a. What are fuzzy techniques?

2b-d. Briefly describe the main steps of fuzzy techniques, and present formulas for these steps:

- eliciting degrees of confidence and forming membership function,
- using “and”- and “or”-operations to estimate the degrees to which different control values are reasonable,
- defuzzification.

3. *Fuzzy techniques: example.* Suppose that we have two rules:

- if a student is tired, the student needs some rest;
- if a student is very tired, the student needs a lot of rest.

A student marked his being tired as 6 on a 0-to-10 scale and being very tired as 4 on this scale. To what extent it is reasonable for this student to rest for an hour? Assume that:

- the degree to which 1 hour means some rest is 0.7, and
- the degree to which 1 hour means a lot of rest is 0.3.

Use min and max as “and”- and “or”-operations.

4. *Which fuzzy techniques?*

- 4a. Which interpolation algorithm should we use when generating a membership function and why?

4b-c. Which “and”- and “or”-operations should we use and why:

- when we control a group of objects, and
- when we control an individual object.

4d. What defuzzification procedure should we use and why?

4e. How fuzzy techniques can be used in explainable AI?

In all these question, just describe the main criterion used for the corresponding selection – e.g., what exactly expression we optimize and why; no need to repeat the derivation of the corresponding formulas from the criterion.

5. *Making deep learning itself more explainable.*

5a. What is a rectified linear activation function and how is it used?

5b. Explain why the rectified linear activation function is empirically the best.

5c. What is pooling and how is it used?

5d. Explain why max- and sum-poolings are empirically the best.

6. *Projects.*

6a. Briefly describe your project.

6b. Briefly describe somebody else's project.