project canvas

Project name Distillation of Ensemble for Uncertainty Quantification for Semantic Classification

Project owner William Krøyer Stentzer
Mads Bjørke Vejbæk

Purpose
What is the intent of this project?
Why are we doing this project?

We're doing this project to test if we can improve upon an open-source LLM's performance

Scope
What does this project contain?
What does this project not contain?

The scope of this project is to do a deep-dive into natural language processing, and the challenges within this field. It does not include inventing a new method for creating LLMs.

Success Criteria
What do we need to achieve in order for the project to be successful?
How can the Success Criteria be measured?

The project is deemed a success if our final distillation shows an improvement, or atleast the same level compared to the base model.

Milestones
When will we start the project and when is the final deadline ?
What are the key milestones and when will they occur?
How can the milestones be measured?

1st handin:
Project description
Project Canvas
Gantt Chart

1st MVP:
Basic model
Basic evaluation

Expanded model with larger ensemble

Expanded model with larger LLM

Finishing the report

Actions
Which activities need to be executed in order to reach a certain milestone?

- Cooperation agreement
- Finding research questions
- Problem definition
- Get approval from supervisors

- Research LLM structure
- PyTorch knowledge
- Dataset analysis and setup
- Coding/1st testing
- Continue report writing, focus on theoretical part

- Evaluation of current models
- Discuss results/outcomes with supervisor
- Continue report writing, focus on results and method.
- Expand code for more robust testing

- Further evaluations on models
- Feedback from supervisor

- Rewrite underwhelming parts of the report
- Make sure there is good continuity through-out the report
- More feedback from supervisor

Outcome
What is the end result?

- A book
- A website
- An event

A report,
Functioning model,
A presentation

Team
Who are the team members?
What are their roles in the project?

William Krøyer Stentzer
Mads Bjørke Vejbæk

Stakeholders
Who has an interest in the success of the project?
In what way are they involved in the project?

Supervisors:
- Michael Riis Andersen
- Morten Mørup

Users
Who will benefit from the outcome of the project?

- Researchers in NLP
- ML engineers
- AI model engineers
- Academics studying NLP techniques

Resources
What resources do we need in the project?
- Physical (office, building, server)
- Financial (money)
- Human (time, knowledge)

- Computational power
- Experts, such as our supervisors
- Online assistance, such as LLMs (ChatGPT) and documentation

Constraints
What are the known limitations of the project?
- Physical (office, building, server)
- Financial (money)
- Human (time, knowledge, politics)

- Computational complexity
- Limited time
- Current lack of knowledge

Risks
Which risks may occur during the project?
How do we treat these risks?

- Model Performance Degradation: To prevent this, frequent testing is required during model distillation
- Overfitting, the model could perform bad on the OOD data: Test on diverse datasets, and cross-validation are posssible solutions for this