COMP492 SENIOR DESIGN PROJECT II PROPOSAL FORM
DEPARTMENT OF COMPUTER ENGINEERING

Project Name

Turkish Question Generation Model

Project Summary (Abstract)

Question Generation is a Deep Learning model that given a paragraph, passage or an entity in Turkish, produces the possible questions that can be answered solely by the given content to the system. Question Generation(QG) Systems are capable of generating various logical questions from the given text input. QG Systems are prevalent in several computer applications such as chatbots, automated grading systems etc.

In this project, our aim is to train different deep learning models for question generation and assess their performances. We find this task worth tackling since there are not many examples of comprehensive studies on the topic and the outcomes of this project will empower the development of more capable-than-ever Question Answering Systems for the Turkish Language.

QG models require extensively large training data, thus, in the last semester we have developed the initial dataset for the development of the QG model. This dataset is the largest Turkish Question Answer dataset currently available. In this semester, we also aim to improve the diversity of our dataset to generate questions in larger extent.

Keywords

Natural Language Processing, Deep Learning, Question Generation

Hardware and Software Requirements

- GPU Optimized Server. Will be used for training deep learning models, storing large amounts of data and as a collaborative environment.

- UNIX Based Environment(s). Given the design and integration of de-facto programming languages, frameworks, tools with Unix based operating systems, along with the flexibility and low-level tools it has to offer, we concluded that UNIX based operating systems will be most suitable for the development process.

- Dataset. Contains Turkish paragraphs, related questions and their answers.

Project Tasks, Time Plan and Deliverables

Table 1: Project tasks, time plan and deliverables

| Task	Start & Due Dates	Deliverable	Evaluation Criteria	Objective
Project Proposal	22/02/2021 05/03/2021	Proposal with clear goals.	Readable, clear	Expressing our intent for the research project and our tasks for this semester.
Literature Review and data collection	01/03/2021 12/03/2021	A collection of related papers and a training dataset for QG.	Moderate size dataset and collection of recent publications	Learning the foundations of the learning-based methods, developing insight about the concepts. Learning state-of-the art in this research area. Collecting the first Turkish QG dataset.
Development Environment Preparation	04/03/2021 10/03/2021	A working environment	Installment of necessary tools, programming languages etc.	Embracing the concepts of the operating system, programming language, frameworks and best practices to be used during the development stages
Revision of the dataset from COMP491	08/01/2021 12/03/2021	Revision of the the dataset	Completeness	Covering and revising the work done in detail from COMP491.
Technical Preparation	12/03/2021 25/03/2021	N/A.	N/A.	Preparing the basis for our further work.
Implementation of Models	19/03/2021 15/04/2021	Working prototypes.	Accuracy, performance, capable models	Generating questions in different forms from a given text by using trained models.
Progress Report & Presentation Preparation	06/04/2021 09/04/2021	Well-prepared report and presentation.	Complete, readable, formatted	Covering the work done in detail up to middle of the semester.
Evaluation of Models	19/04/2021 30/04/2021	Report for evaluation of models	Comprehensive, readable, clear	Evaluating the trained models with the state-of-art QG evaluation metrics
User Interface Development	03/05/2021 21/05/2021	Simple UI for users to interact	Simple to use, error-free usage	Enabling the users to interact with the QG model
Final Report & Presentation Preparation	20/05/2021 02/06/2021	Well-prepared report and presentation	Complete, readable, formatted	Covering the work done in detail.
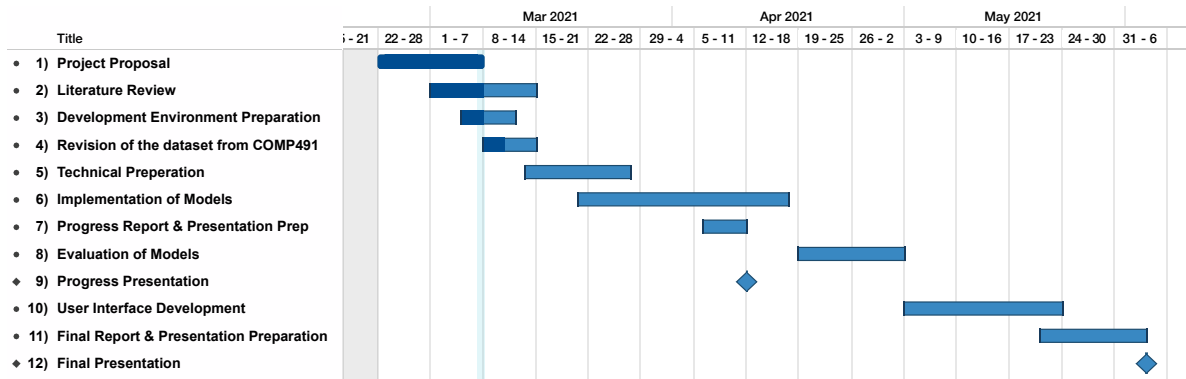
Figure 1: Gantt chart of the project

Project Team and Authority Information

Table 2: Project team and authority information

Proposal Date	08/03/2021
Academic Term of Project Delivery	2020-2021, Spring
Project Team Members	Alp Gokcek, #041701014, Computer Engineering Erdal Sidal Dogan, #041701076, Computer Engineering
Advisor(s)	Seniz Demir