**“Kyrgyz Headline Generation” Project Report**

**Done by: Aktan Raimbekov, Dastan David uulu, Elmirbek Orozobekov, Bayizbek Kutmanov, Argen Shamyrov, Elzar Zholdoshev**

**Main goal of the project:**

We wanted to create an ai, fine-tuned especially for Kyrgyz language Headline Generation based on Text Summarization. The ai should make short Headline using long text inserted.

**Team Contribution:**

Aktan Rayimbekov - Main coder

Dastan David uulu - Coders helper

Elmirbek Orozobekov - Database collector

Elzar Zholdozhev - Presentation design Creator

Argen Shamyrov - Presentation text Creator

Bayizbek Kutmanov - Presentator

**What we learned?**

We learned that Multilingual models based on transformers can be fine-tuned for underrepresented languages.

Preprocessing and tokenizing are very important for model’s good performance

Also we learned that even small datasets if they are made with proper techniques can yield good results.

**What can be improved:**

We could increase our dataset and improve the technique it’s made of so the model made better headlines.

Also would be better if we didn’t use multilingual models pretrained on much languages but used models specially trained for Kyrgyz language or if we trained it by ourselves so the model would understand the text better and generate better headlines