Group Name: individually
Name: Kristina Kaliagina
Email: kalyagina.kristina@gmail.com
Country: Russia
College/Company: Graduated from the University “Higher School of Economics”
Specialization: NLP

Deliverables

Problem description:

The task of the project is to classify tweets from Twitter, namely, it is necessary to create a model that will help determine whether a particular tweet belongs to such a type of speech as hate speech. Hate Speech, in simple terms, is offensive language directed at individuals or groups based on their affiliation, interests, and characteristics, such as their religion, nationality, race, color, origin, gender, or other identity factor.

The task is quite difficult because of the inherent complexity of natural language constructs - different forms of hatred, different types of goals, different ways of representing the same meaning.

Business understanding:

Social media has become the main driver of social change in global society. The consequences of events taking place in one corner of the world are reflected around the globe in different regions. This is because the vast amount of data generated on these platforms reaches the far corners of the world in the blink of an eye. The developers of these platforms face numerous challenges to make cyberspace as inclusive and healthy as possible. However, in recent years, the phenomena of offensive speech and hate speech have been spreading with greater force. Despite manual efforts, the scale of this problem is so huge that it cannot be solved with coordinated teams. In fact, an automated technique needs to be developed that detects and removes offensive and hateful comments before their harmful effects materialize.

The detection of such hate speech is important for the analysis of public sentiment. User groups in relation to another group, as well as to prevent illegal actions. It's also useful to filter tweets before content recommendations or explore AI chatbots on tweets.

Project lifecycle along with deadline