

Моделовање моралних и емоционалних компоненти језика у класификацији конверзационих текстова

Кандидат: Милена Шошић

Конверзационе текстуалне поруке представљају један од главних начина комуницирања у савременом друштву. Са развојем информационих технологија у претходној деценији, појавили су се бројни алати за комуницирање као што су поруке електронске поште, поруке као објаве и коментари на друштвеним мрежама, поруке у алатима за инстант комуницирање и поруке за комуницирање са алатима за аутоматско генерисање одговора. Сви ови типови порука имају своју структуру и могу се класификовати, било појединачно или као скупови порука који припадају једном конверзационом низу, у различите врсте категорија.

Област рачунарске лингвистике је за неке светске језике већ постигла одређени напредак у откривању моралних и емоционалних компоненти у тексту, али су потребна додатна истраживања да би се у потпуности разумели ови сложени и апстрактни концепти, као и њихова међусобна повезаност. За српски језик овакви ресурси још увек нису развијени.

Класификовање конверзација коришћењем карактеристика моралних и емоционалних компоненти језика представља задатак који још увек није довољно истраживан. Предмет овог истраживања биће анализа утицаја моралних и емоционалних компоненти језика у задацима класификације конверзационих текстова. Утицај ових компоненти ће се мерити по степену њиховог доприноса и значаја на успешност класификације односно на успешност поделе података у категорије дефинисане задатком класификације. Применом математички заснованих метода за проналажење значајних атрибута класификације утврдиће се у којој мери морални и емоционални атрибути узимају учешће у овако утврђеним скуповима. У оквиру класификације конверзационих текстова на енглеском језику биће предложен нови, хибридни приступ који користи методе машинског учења за класификацију ове врсте података. У оквиру истраживања биће развијени нови ресурси и предложене технике које би омогућиле примену предложеног приступа на конверзационе текстове написане на савременом српском језику.