

STUDENTŲ ĮTRAUKIMO Į MOKSLINĘ VEIKLĄ SKATININAMOJO KONKURSO TEMA

Temos pavadinimas: Kontekstinės informacijos panaudojimas lietuvių kalbos vaizdų aprašymo sistemose

Tikslas: Ištirti papildomos tekstinės informacijos panaudojimo galimybes kuriant tikslesnius vaizdų aprašymus lietuvių kalba.

Trumpas temos vykdymo aprašymas (ne daugiau kaip 2000 ženklų):

Studentas atliks vaizdų aprašymo (angl. *image captioning*) sistemų, multimodalinių modelių ir papildomos kontekstinės informacijos panaudojimo metodų analizę. Tyrimo metu bus nagrinėjama, kaip papildoma tekstinė informacija, pavyzdžiui, straipsnių antraštės, naujienų tekstai gali pagerinti automatiškai generuojamų vaizdų aprašymų kokybę lietuvių kalba.

Studentas parengs nedidelį eksperimentinį duomenų rinkinį, sudarytą iš vaizdų ir juos papildančios tekstinės informacijos. Naudojant šiuolaikinius vaizdo ir kalbos modelius (pvz., Qwen-VL, Gemma, Intern ar kitus atvirojo kodo modelius) bus atliekami eksperimentai, lyginant aprašymus, generuojamus tik iš vaizdo ir generuojamus naudojant papildomą kontekstinę informaciją.

Tyrimo metu bus vertinamas sugeneruotų aprašymų tikslumas, išsamumas, kalbinė kokybė, kontekstinis tinkamumas ir gebėjimas pateikti informaciją, kurios neįmanoma tiksliai nustatyti vien iš vaizdo turinio. Taip pat bus analizuojamos dažniausiai pasitaikančios klaidos ir netikslumai. Gauti rezultatai bus lyginami naudojant automatines vertinimo metrikas bei kokybinę analizę.

Atliekant šį tyrimą reikės:

1. Pasirinkti vaizdų duomenų aibę ir papildomus tekstinės informacijos šaltinius bei parengti eksperimentinį duomenų rinkinį.
1. Pasirinkti ir išbandyti vieną ar kelis vaizdo ir kalbos modelius, generuojant vaizdų aprašymus naudojant tik vaizdą ir naudojant papildomą kontekstinę informaciją.
2. Įvertinti sugeneruotų aprašymų kokybę pagal tikslumą, informatyvumą, kalbinę kokybę ir atitikimą pateiktam kontekstui.
3. Atlikti palyginamąją rezultatų analizę, įvertinti kontekstinės informacijos įtaką lietuviškų vaizdų aprašymų kokybei ir aprašyti tyrimo rezultatus.

Temą siūlantis mokslininkas/dėstytojas:

TOPIC OF A COMPETITION PROMOTING STUDENT ENGAGEMENT IN SCIENTIFIC ACTIVITIES

Topic::
Goal:
Short description (max. 2000 characters):
Supervisor researcher/lecturer: