**Ajaseoste märgenduse masinõppimine**

Juhendaja: Siim Orasmaa ( *siim . orasmaa @ ut . ee )*

Ajasemantilise märgenduse eesmärgiks on tuua formaalsel kujul esile loomulik keele tekstis peituvad ajaseosed, võimaldades tekstis viidatud sündmuste omavahelist järjestamist ning ajateljele paigutamist. Selline märgendus toetab mitmeid keeletehnoloogia lõpprakendusi, nt automaatset kronoloogiate genereerimist, küsimustele vastamist ning sisukokkuvõtete tegemist.

Antud töö eesmärgiks on luua prototüüp süsteemist, mis tuvastab automaatselt eestikeelsest tekstist ajaseosed. Selleks tuleb kasutada juhendatud masinõppe lähenemist: katsetada Eesti TimeML korpusel erinevaid masinõppealgoritme ja tunnuste kombinatsioone, eesmärgiga leida parimad mudelid ülesande lahendamiseks. Programmeerimiskeelena tuleks kasutada Pythonit ning soovitavalt teegi *scikit-learn* ( http://scikit-learn.org ) vahendeid.

Viited:

Eesti TimeML korpus:
http://keeleressursid.ee/et/keeleressursid-cl-ut/korpused/83-sample-data-articles/clutee-lehed/276-ajasemantiliste-margendustega-tekstikorpus

Pustejovsky, J., & Verhagen, M. (2009). SemEval-2010 task 13: evaluating events, time expressions, and temporal relations (TempEval-2). In Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions (pp. 112-116). Association for Computational Linguistics.

Verhagen, M., Sauri, R., Caselli, T., & Pustejovsky, J. (2010). SemEval-2010 task 13: TempEval-2. In Proceedings of the 5th international workshop on semantic evaluation (pp. 57-62). Association for Computational Linguistics.