

Poročilo 3. domače naloge pri predmetu Uvod v odkrivanje znanj iz podatkov

Uroš Vaupotič 64140387

November 27, 2017

1 Uvod

Cilj naloge je bil čim bolj natančno napovedovanje časa prihodov avtobusov glede na čas odhodov.

2 Ocenjevanje točnosti

Interno točnosti sem preverjal na podatkih za učenje. Implementiral sem standardno formulo za Mean absolute error (MAE) in računal samo za podatke v mesecu novembru, saj je ta še najbolj podoben decembru.

3 Napovedni modeli

Vsi meseci Vzel sem vse mesece v učnih podatkih

Prazniki Med učne podatke sem dodal praznike in dela proste dneve tako, da so podatki vsebovali podatek, ali je vožnja potekala na praznik.

Nočna ali dnevna vožnja K učnim podatkom sem dodal podatek, ali se vožnja dogaja med 6. in 20. uro, saj sem predpostavil, da je takrat več prometa, ki bo upočasnil vožnjo.

Polovica ure Dodal sem podatek, ki je učne podatke razdelil na vožnje, ki so se začele v prvi polovici, oziroma v drugi polovici ure. Predpostavil sem, da se pri večini dejavnosti povezanih s prometom aktivnosti začnejo ob polni uri, torej bo imel avtobus, ki je z vožnjo začel v drugi polovici ure, večjo možnost zgoščenega prometa.

4 Rezultati

Vsaka metoda vsebuje tudi metode v tabeli nad sabo.

Table 1: Rezultati

IME METODE	OCENA	KOMENTAR
Vsi meseci	152.32	Precej bolši rezultat od average
Prazniki	146.91	Rezultat se je občutno izboljšal
Nočna ali dnevna vožnja	146.84	Majhna izboljšava
Polovica ure*	146.84	Skoraj zanemarljiva izboljšava