Radili smo novi nacin odabira značajki. Kako su značajke kategoričke možemo napraviti chi2 test nezavisnosti oznake o svakoj značajki. Chi2 test računa koliko se očekivana i izračunata frekvencija događaja razlikuju (https://towardsdatascience.com/chi-square-test-for-feature-selection-in-machine-learning-206b1f0b8223). Kako je to statistički test, možemo odbacivati nultu hipotezu s nekom razinom signifikantnosti. U ovom slučaju, nulta hipoteza je da je oznaka neovisna o značajki. Mi smo u ovom radu razinu signifikantnosti stavili na 0.01 (1%). Nakon što smo napravili test za svaku značajki uzeli smo samo one značajke koje bi odbacile nultu hipotezu. Takvim pristupom vidimo da naivni Bayesov klasifikator pokazuje najveće povećanje metrika. U slučaju ostalih klasifikatora, ova metoda odabira značajki daje odprilike iste rezultate kao i ostale metode.