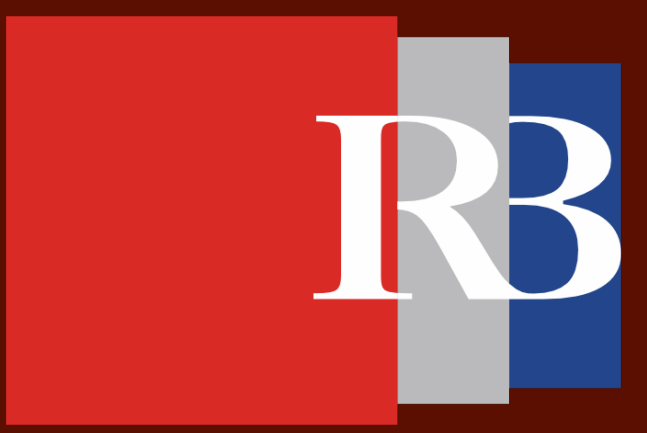


KOREKCIJA PRISTRANOSTI U GENOMSKIM PODACIMA POMOĆU DUBOKOG UČENJA

Ira Renko¹, Damir Baranašić¹

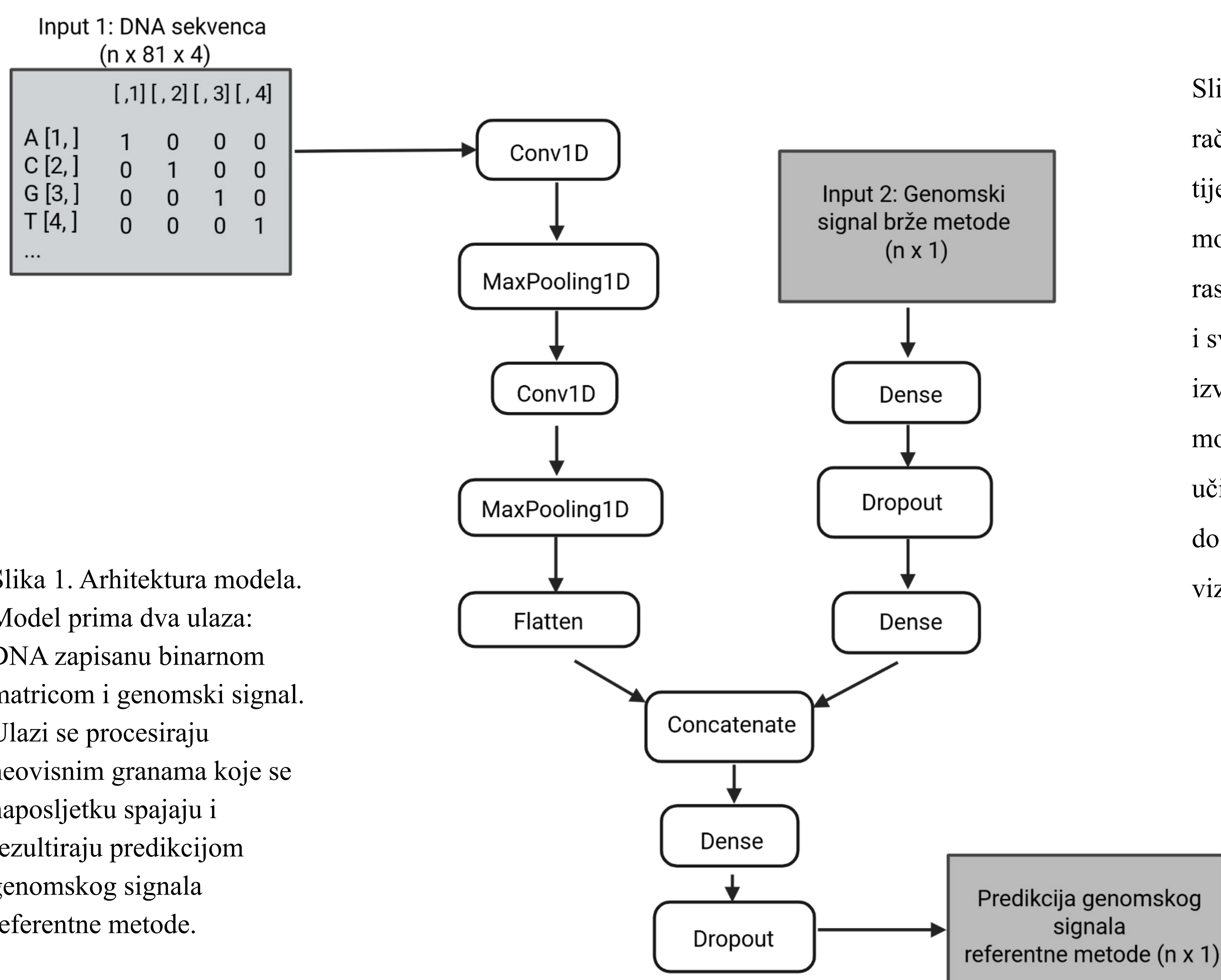


1. Institut Ruđer Bošković, Zavod za elektroniku, Laboratorij za računalnu biologiju i translacijsku medicinu

UVOD

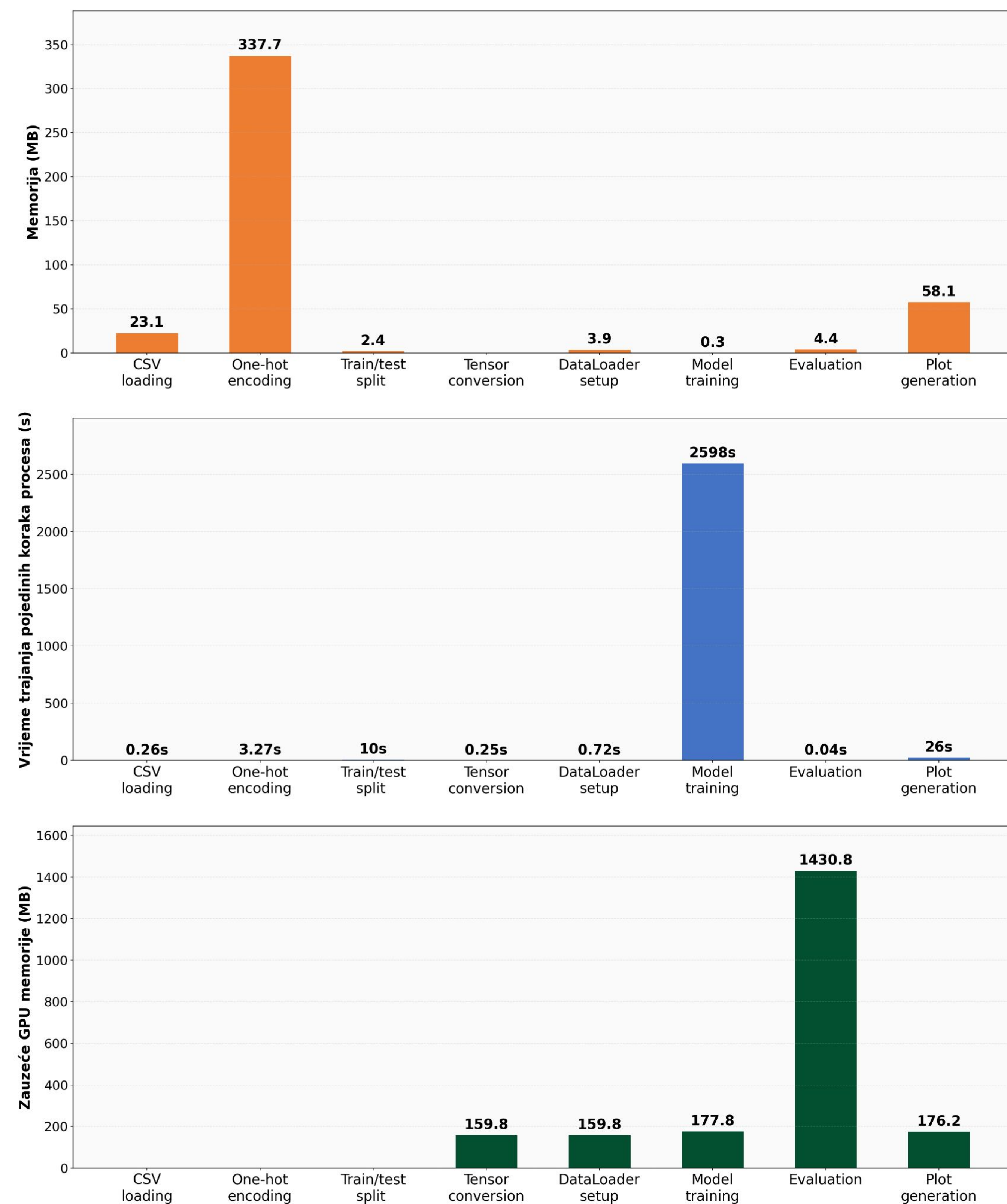
Genomska mjerenja generiraju ogromne količine podataka, milijune očitavanja po eksperimentu. Neke metode su točnije, ali skuplje, kompleksnije i nisu uvijek izvedive; druge rade s manje početnog materijala, brže i jeftinije, ali su manje pouzdane. Koristimo duboko učenje za korekciju sustavnih pristranosti jednostavnijih metoda koristeći točnije metode kao reference.

METODA



Slika 2. Potrošnja računalnih resursa tijekom korištenja modela. Prikazana je raspodjela memorije i sveukupno trajanje izvođenja koraka modela od učitavanja podataka do evaluacije i vizualizacije.

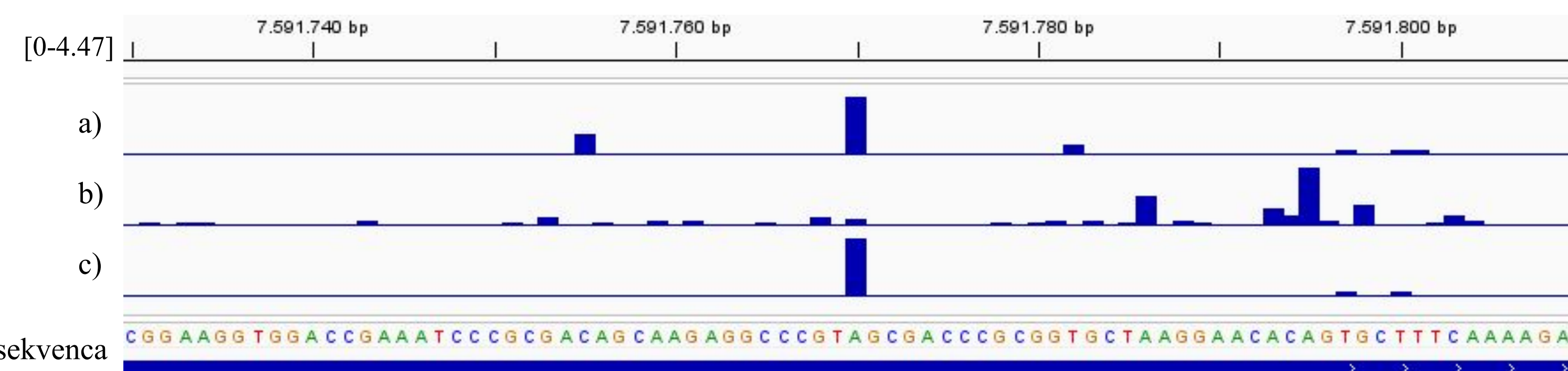
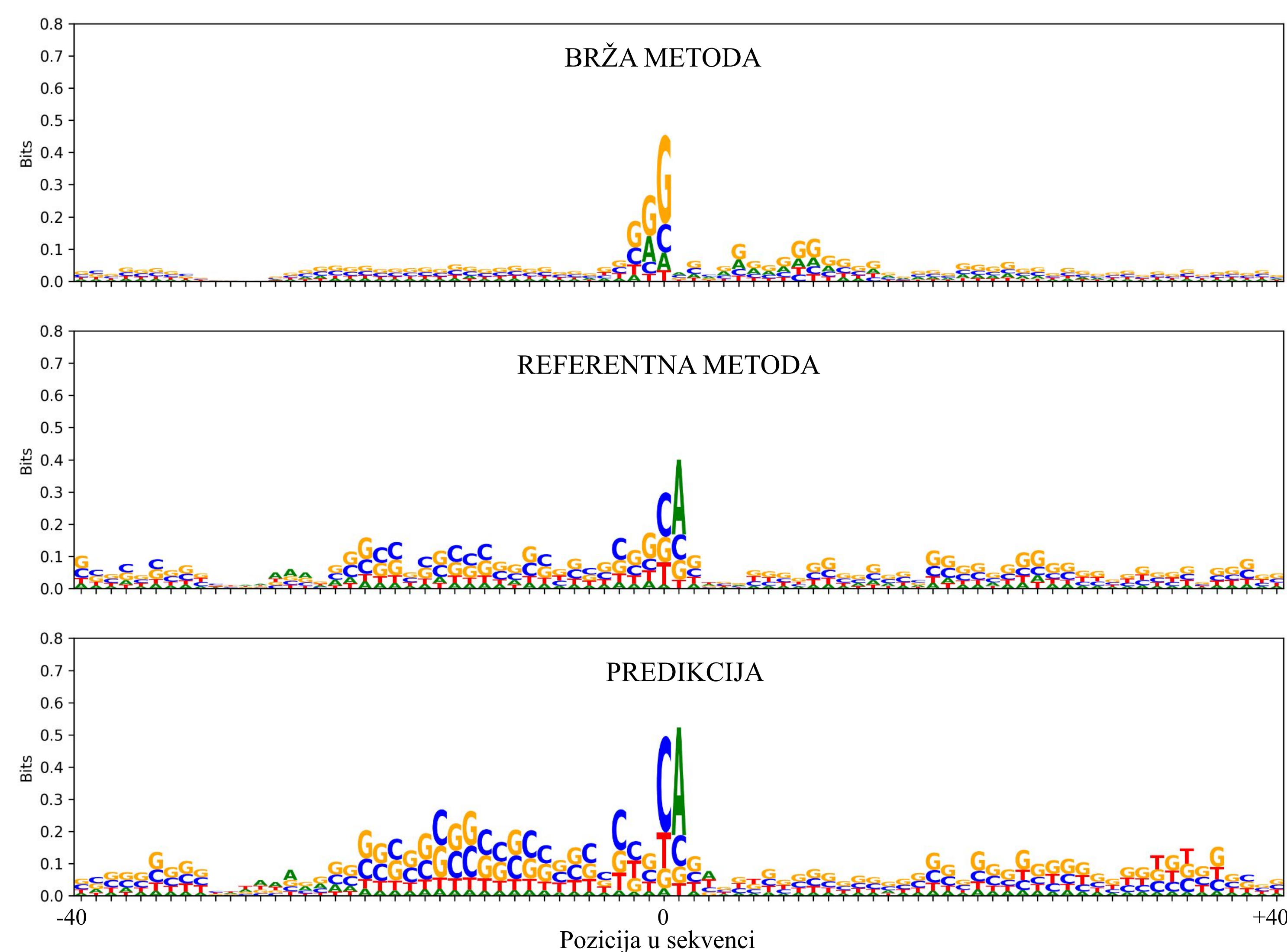
REZULTATI



Slika 3. Prikaz rezultata predikcije genomskog signala referentne metode.

Usporedbom izgleda genomskog signala brže, referentne metode i izlaza modela vidljivo je kako model uspješno koregira sustavnu pristranost jednostavnije metode.

REZULTATI



Slika 4. Prikaz genomskog signala a) brže metode, b) referentne metode i c) predikcije u genomskom pregledniku. Vizualizacijom u pregledniku rezultati predikcije su dodatno validirani.

ZAKLJUČAK

Modeli trenirani na HPC sustavu s CUDA akceleracijom omogućili su predikciju za 128 261 uzorka. Korekcija sustavne pristranosti uspješno je validirana usporedbom s referentnom metodom.

[1] Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P., & Lenhard, B. (2015). CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic acids research* 43(8), e51-e51.
 [2] Poulain, S., Kato, S., Arnaud, O., Morlighem, J. É., Suzuki, M., Plessy, C., & Harbers, M. (2017). NanoCAGE: a method for the analysis of coding and noncoding 5'-capped transcriptomes. In *Promoter Associated RNA* (pp. 57-109). Humana Press, New York, NY.
 [3] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroury, A., Shuai, B. i sur. (2018). Recent advances in convolutional neural networks. *Pattern recognition*, 77, 354-377.



Funded by the European Union NextGenerationEU