

Report za zimný semester

Zoznámenie sa s problematikou. Zoznámenie sa s formátom hdf5, a s knižnicami, ktoré s týmto formátom pracujú. Takisto špecifikácie formátu fast5(hdf5) a fastq využívaných pri projekte. Implementácia naivného (algoritmicky nezaujímavého) riešenia, ktoré načíta potrebné dáta z pôvodného fast5 súboru, a fastq súboru ktorý vznikol ako výstup guppy basecallera. Následne program priradí jednotlivé reads z oboch súborov k sebe (podľa idčka), a vytvorí pre ne tabuľku Move, ktorú zapíše do fast5 súboru rovnako, ako to robili predošlé verzie basecallera guppy.

Plán na letný semester:

Návrh a implementácia heuristickej metódy na výpočet Move, najprv v jazyku Python, neskôr prechod na rýchlejší jazyk (c++). V poslednej fáze Dokumentácia projektu.