

## PyFlowSOM: Project Voorstelling

### 1. Samenvattende beschrijving

We implementeren en optimaliseren het FlowSOM algoritme en het toepassen ervan voor grote datasets in python.

FlowSOM staat voor Flow cytometry Self-Organizing Map. Het wordt gebruikt om cellen te clusteren op basis van bepaalde eigenschappen. Dit is waardevol omdat het onderzoekers toelaat om hoogdimensionale flowcytometrie gegevens te bestuderen.

De computationele biologie gebeurt steeds vaker in python. We zetten daarom het R package om naar python. Op deze manier verhogen we de toegankelijkheid voor onderzoekers en creëren we de mogelijkheid tot integratie met nieuwe en bestaande systemen. Daarnaast worden de datasets steeds groter waardoor er een optimalisatie nodig is. Het doel is om deze te leveren.

Naast het implementeren van de library zullen we de ook de implementatie onderwerpen aan metingen op vlak van tijd en geheugengebruik. Tot slot verbeteren we de UI/UX van het package.

### 2. Ontwikkelings Roadmap

De roadmap voor de ontwikkeling van PyFlowSOM ziet er als volgt uit.

Startdatum	Einddatum	Item	Label
May 9, 2023	May 10, 2023	Opgave doornemen	Onderzoek
May 11, 2023	May 11, 2023	Project voorstelling schrijven	Raporteren
May 12, 2023	May 12, 2023	Linters toevoegen	Developer Tools
May 12, 2023	May 12, 2023	Test framework toevoegen	Developer Tools
May 12, 2023	May 14, 2023	Bestaande R package analyseren	Onderzoek
May 14, 2023	May 14, 2023	API uitdenken	Onderzoek
May 14, 2023	May 14, 2023	Optimalisatie methodes uitdenken	Onderzoek
May 12, 2023	May 14, 2023	Relevante libraries opzoeken	Onderzoek
May 14, 2023	May 18, 2023	Eerste FlowSOM implementatie	Implementeren
May 14, 2023	May 15, 2023	1. Data preprocessing	Implementeren
May 15, 2023	May 16, 2023	2. SOM training	Implementeren
May 16, 2023	May 17, 2023	3. Meta clustering	Implementeren
May 17, 2023	May 18, 2023	4. Quality control	Implementeren
May 18, 2023	May 18, 2023	5. Basic visualization	Implementeren
May 19, 2023	May 24, 2023	Optimalisatie	Optimalisatie
May 19, 2023	May 24, 2023	Benchmarking	Optimalisatie
May 25, 2023	May 25, 2023	Stroomlijnen van de codebase	Optimalisatie
May 26, 2023	May 27, 2023	Uitgebreid testen	Testen
May 28, 2023	May 29, 2023	Verslag schrijven	Raporteren
Jun 14, 2023	Jun 14, 2023	Mondelinge verdediging voorbereiden	Raporteren
Jun 15, 2023	Jun 15, 2023	Verdediging	Raporteren

#### *Extra quality of life toevoegingen*

May 23, 2023	May 25, 2023	Meer gepolijste visualisaties	Implementeren
May 23, 2023	May 25, 2023	Externe documentatie voorzien	Developer Tools
May 23, 2023	May 25, 2023	Ondersteuning voor meer data formaten	Implementeren
May 23, 2023	May 25, 2023	Geavanceerdere foutafhandeling	Implementeren