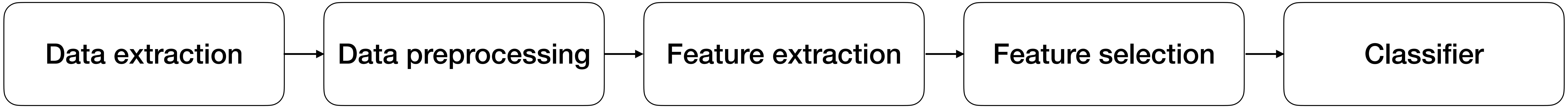
Recognition of activity from left & right hand using EEG data

Bio-Inspired Computers

Sabína Gulčíková, xgulci00

Familiarize yourself with brain EEG (electroencephalogram) and data provided by the supervisor. The EEG Dataset for Motor Imagery BCI (http://gigadb.org/dataset/100295) provides EEG recordings Motor Imagery (MI) Brain Computer Interface (BCI) from 52 subjects. In addition to that, psychological and physiological questionnaires, EMG datasets, location of 3D EEG electrodes and EEGs for non-task related states is also available in the dataset. There are two events in the data, that is, left hand movement and right hand movement. **Investigate, select and design an evolutionary computation algorithm to classify the data into these 2 classes.** Explore and propose the hardware implementation of the algorithm.

Implementation



Implementation

Data extraction	Data preprocessing	Feature extraction	Feature selection	Classifier
	Relevant channels extraction	FeatureExtractor	Evolutionary algorithm	Support Vector Machine
	Baseline correction	MNE-Features library		scikit-learn library
	Spectral filtration (Butterworth)			

Implementation

Data extraction → Data preprocessing → Feature extraction → **Feature selection** → Classifier

Data preprocessing
- Relevant channels extraction
- Baseline correction
- Spectral filtration (Butterworth)

Feature extraction
- FeatureExtractor
- MNE-Features library

Feature selection
- **Evolutionary algorithm**

Classifier
- Support Vector Machine
- scikit-learn library

Features

- Statistical
  - mean
  - variance
  - skewness
  - …
- Frequency Domain
  - spectral entropy
  - energy in frequency bands
  - …
- Complexity Measures
  - Hurst exponent
  - approximate entropy
  - …
- Wavelet Transform Features
- Nonlinear Interdependence Measures
- …

Experiments design

- Hypothesis: the more features the better.
- Goal is to find the optimal combination of features, and experiment with different settings of parameters of evolution.

Experiments design

- Individual is a tuple x, where $x \in \{0,1\}^n$, and n is the number of available features.
- `VALID_FEATURES` is a list of possible features.

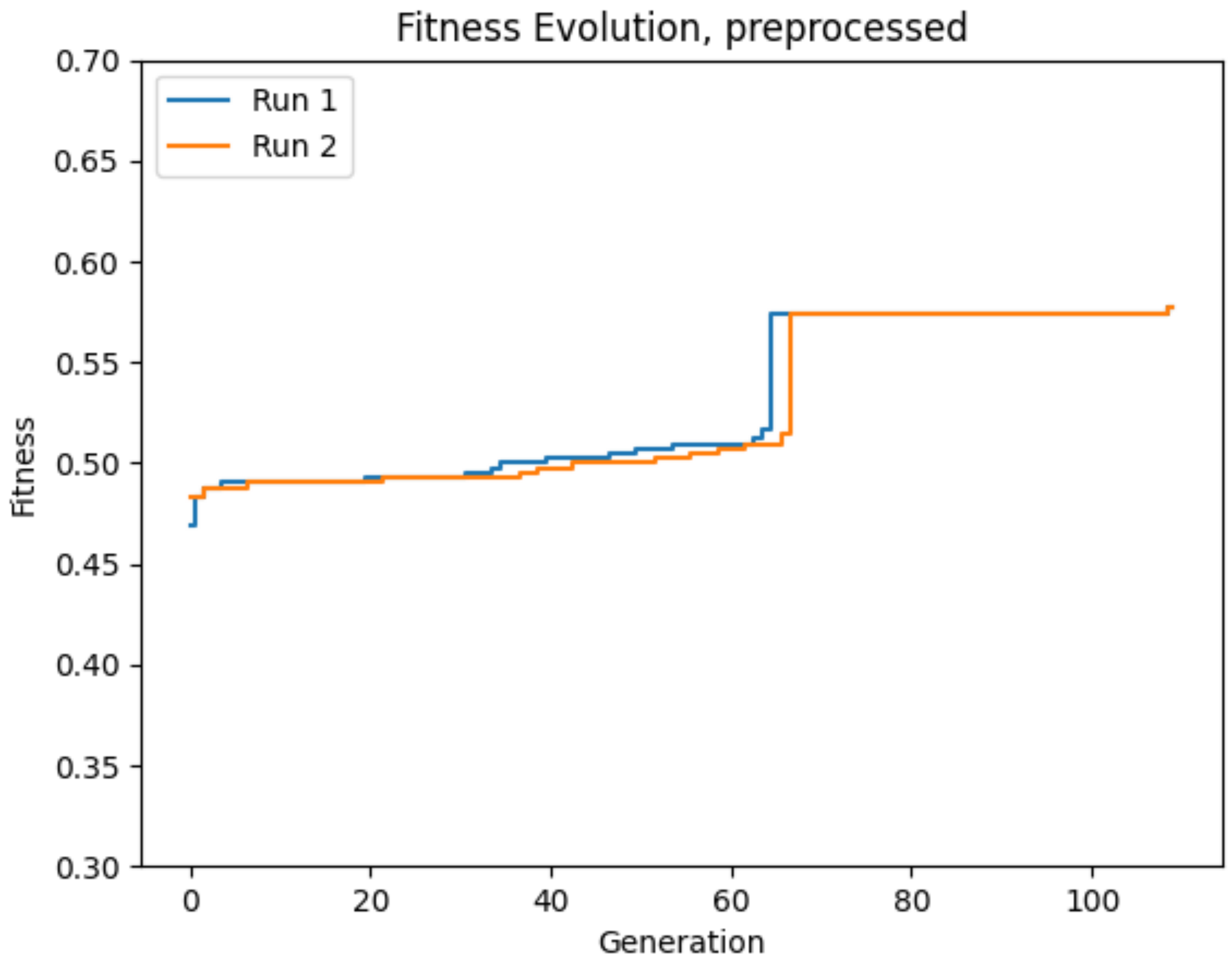
```
individual1 =
[     0       ,      1     ,          1          ,…,       1        ,        0        ]
VALID_FEATURES =
['app_entropy','higuchi_fd','hjorth_complexity',…,'zero_crossings', ‘phase_lock_val']
selected_feats =
[‘higuchi_fd','hjorth_complexity',…,'zero_crossings']
```

Experiments design

- Fitness = {accuracy, weighted sum of accuracy and F1 score}
- Population initialization = {random, single feature vectors + random}
- Preprocessing = {included, not included}

Results

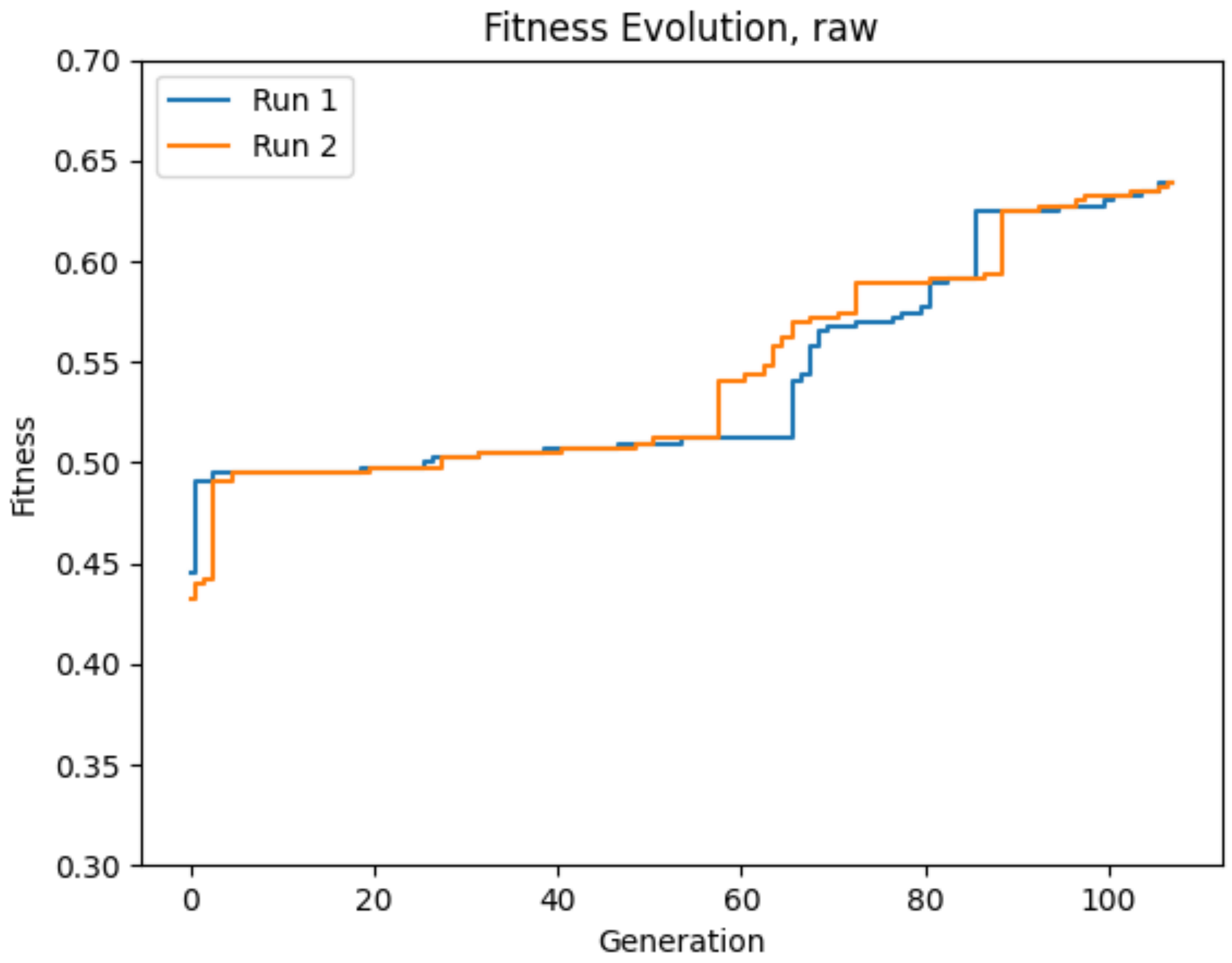
fitness = accuracy



Best individual's fitness: 0.57451,

Best individual:

[0, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0]



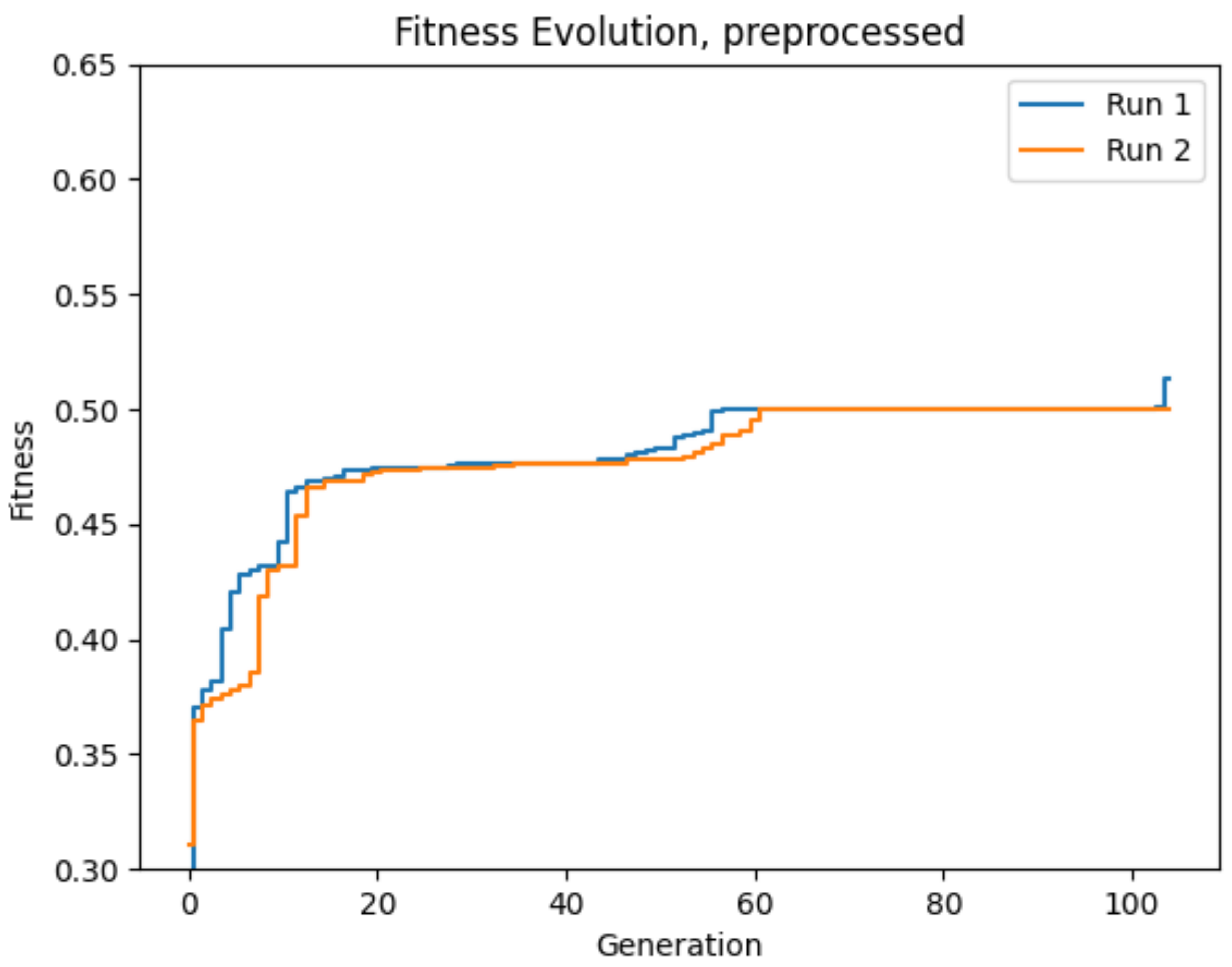
Best individual's fitness: 0.64182,

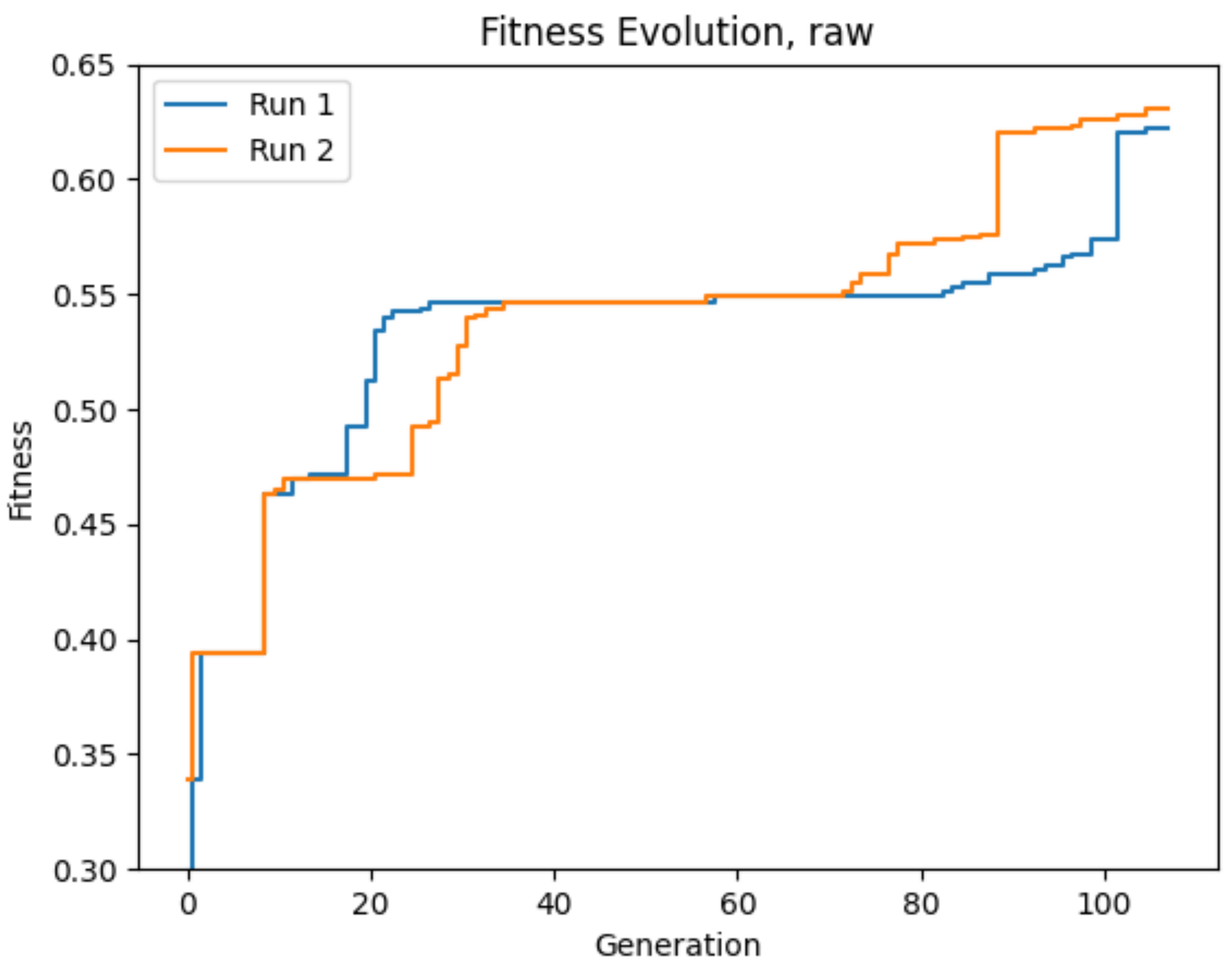
Best individual:

[0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0]

Results

$$\textbf{fitness} = w_a \cdot \textbf{accuracy} + w_f \cdot \textbf{F1 score}$$





Best individual's fitness: 0.57451,

Best individual:

[0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1]

Best individual's fitness: 0.63101,

Best individual:

[0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0]

Conclusion

- Room for improvements:
  - Deeper domain knowledge + proper preprocessing
  - Make better use of the dataset for defining more precise fitness function
  - Make use of *niche evolution* to gain better control of what features can be combined