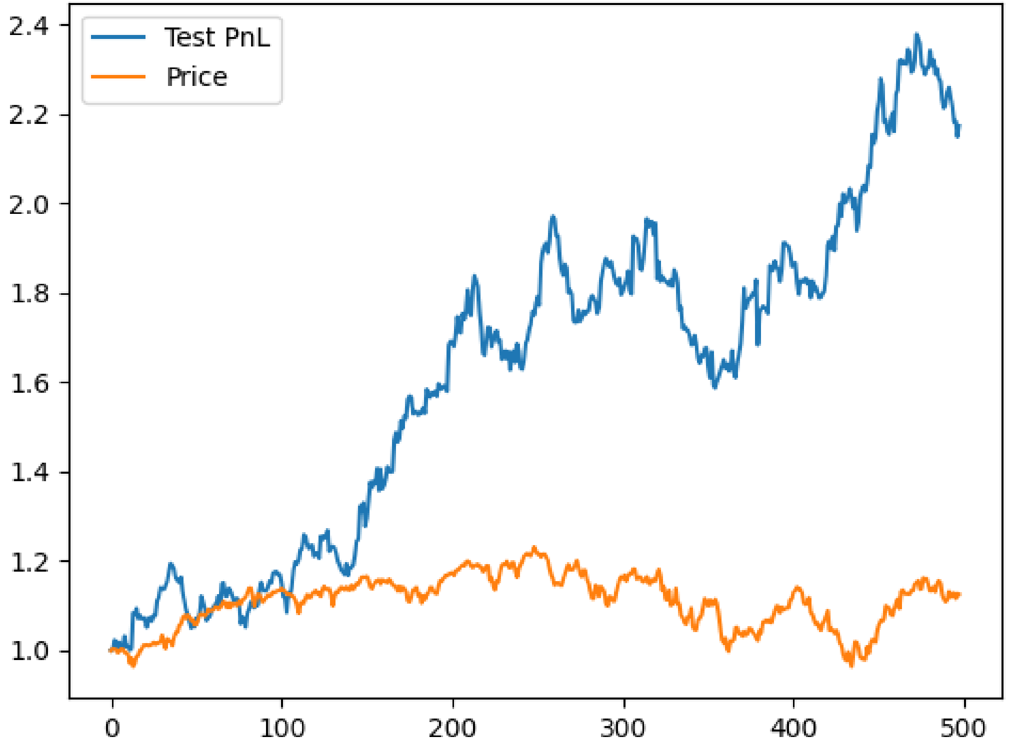
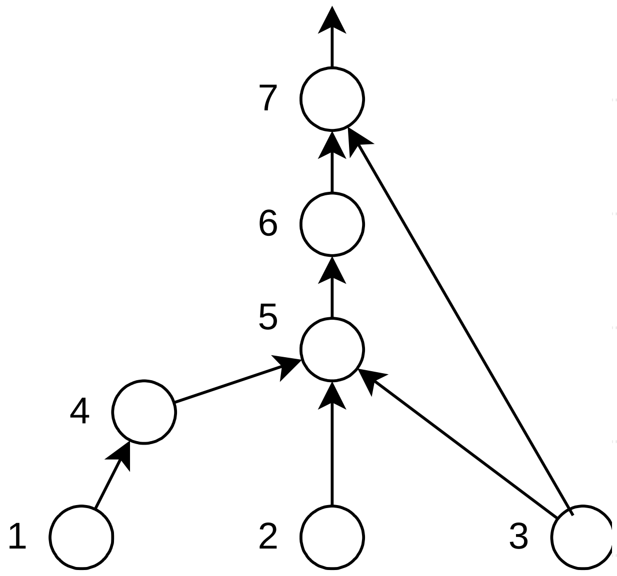
An Alternative Approach



An Alternative Approach

Consider that:

- Dynamic neural network structure → set of dimensions (freedom) of the objective space



Weight W

$= w_{(1,4)}+w_{(2,5)\,)}+w_{(3,5)}+w_{(4,5),\,)}+w_{(5,6)}+w_{(6,7)}+w_{(3,7)}$

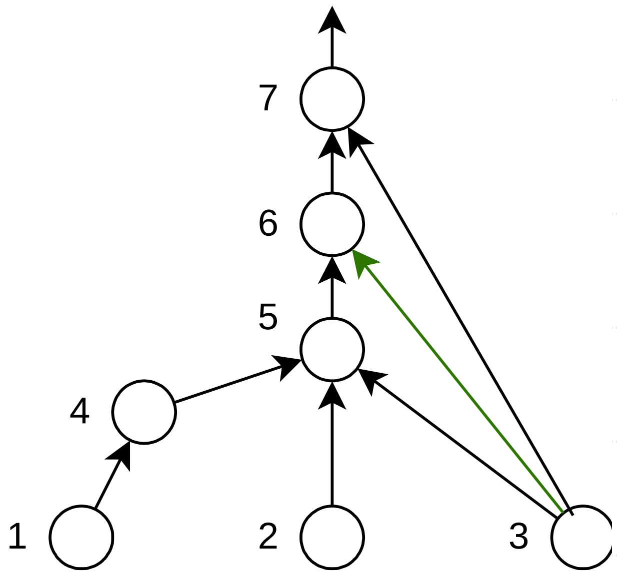
Dimensions of objective space:

{(1,4), (2,5), (3,5), (4,5), (5,6), (6,7), (3,7)}

An Alternative Approach

Consider that:

- Dynamic neural network structure → set of dimensions (freedom) of the objective space



Weight W

$= w_{(1,4)}+w_{(2,5)\,)}+w_{(3,5)}+w_{(4,5),\,)}+w_{(5,6)}+w_{(6,7)}+w_{(3,7)}+w_{(3,6)}$

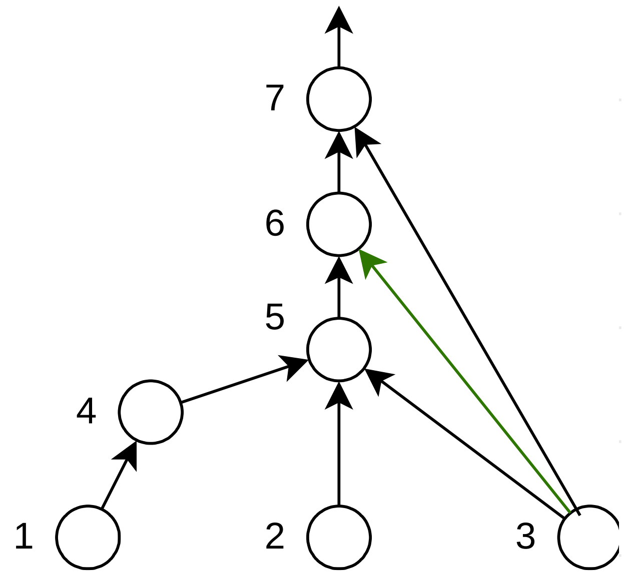
Dimensions of objective space:

{(1,4), (2,5), (3,5), (4,5), (5,6), (6,7), (3,7), (3,6)}

An Alternative Approach

Consider that:

- If the new dimension is introduced gradually from 0, the change in the whole objective space is gradual



$$\forall\, b > 0, \exists\, c \text{ s.t. } f(W|(3{,}6) \notin D) - f\left(W\middle||w_{(3,6)}| < c\right) < b \;\; \forall\, W$$

Weight W

= $w_{(1,4)}+w_{(2,5)\,)}+w_{(3,5)}+w_{(4,5),\,)}+w_{(5,6)}+w_{(6,7)}+w_{(3,7)}+w_{(3,6)}$

Dimension set D of objective space:

{(1,4), (2,5), (3,5), (4,5), (5,6), (6,7), (3,7), (3,6)}

An Alternative Approach

Consider that:

- Choice of dimension set D is discrete, but the effect of adding new dimensions on the objective space is continuous for any W of any original D
- Discrete set D, is not differentiable
- Its effects to the output space and evaluation space are continuous
- GA can be adopted to optimize the choice of D

An Alternative Approach

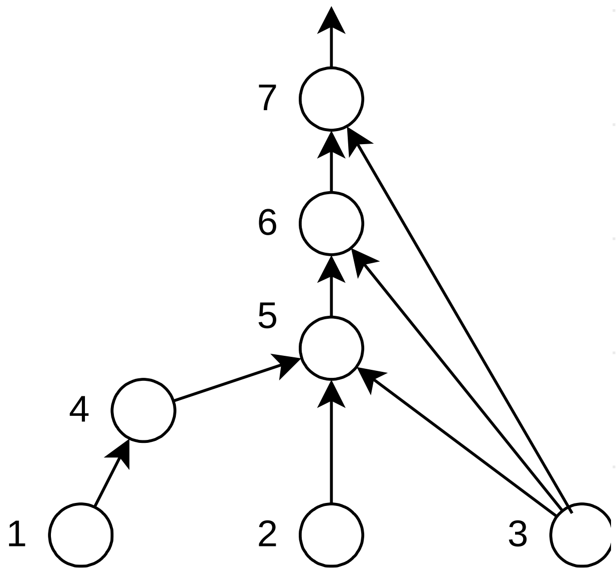
Consider that:

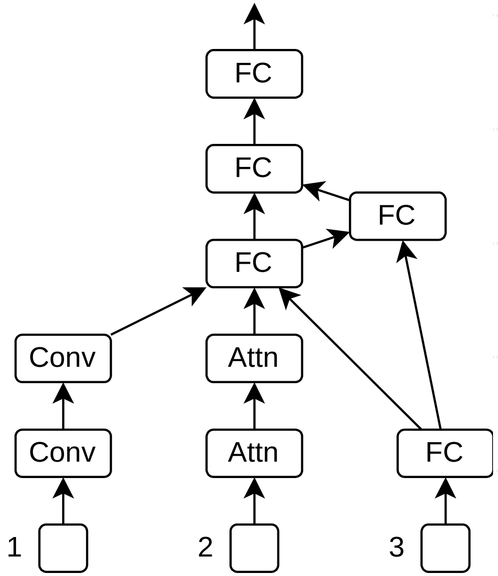
- Given a dimension set D, an objective space $f_D: \mathbb{R}^{|D|} \rightarrow \mathbb{R}$ can be constructed
- GD can be applied to optimize W on f_D
- After certain iteration, W would enter local minimum in f_D
- Add new dimension d to D to get $\widetilde{D}$ and construct $f_{\widetilde{D}}$, where $f_{\widetilde{D}}(W) = f_D(W)\ \forall\ W$ as $W_d = 0$
- Now, W can search also in dimension d by removing the constraint $W_d = 0$
- W can escape the local minimum in f_D

An Alternative Approach

Consider that:

- NEAT tries to do so, but it is extremely slow and ineffective as it adds d one by one
- → Modular NEAT, add a set of dimensions at once





Modular NEAT

- A model is treated as an acyclic directed graph of modules
- Each node represents a module
- Each edge directs the output of a module to the input of another module
- To allow new edges, data shapes and spatial property have to be specified
- Genome: structure (modules and edges) and weights
- Genotype → Phenotype
- Evolves via natural selection and mutation

Genome

- NEAT Genome Properties
- **Node**: Node index, node type
- **Edge**: In node, out node, weight, availability, innovation
-
- MNEAT Genome Properties
- **Node**: Node index, module, input shape, output shape, input spatial property, weights
- **Edge**: In node, out node, weight

Phenotype

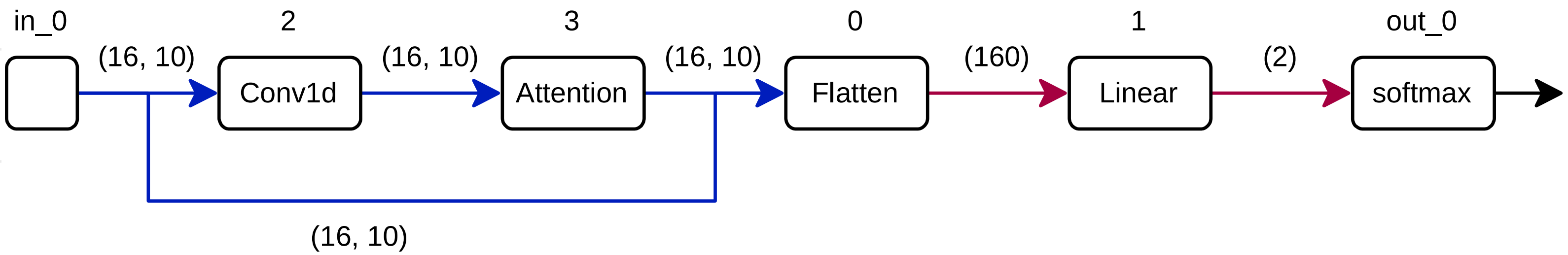
- Node Genes:

index: in_0	**index**: 0	**index**: 1	**index**: 2	**index**: 3	**index**: out_0
module: None	**module**: Flatten	**module**: Linear	**module**: Conv1d	**module**: Attn	**module**: Softmax
in shape: None	**in shape**: 16, 10	**in shape**: 160	**in shape**: 16, 10	**in shape**: 16, 10	**in shape**: 2
out shape: 16, 10	**out shape**: 160	**out shape**: 2	**out shape**: 16, 10	**out shape**: 16, 10	**out shape**: None
out spatial: True	**out spatial**: False	**out spatial**: True	**out spatial**: True	**out spatial**: True	**out spatial**: False

- Edge Genes:

in_0 → 0	0 → 1	1 → out_0	in_0 → 2	2 → 3	3 → 0

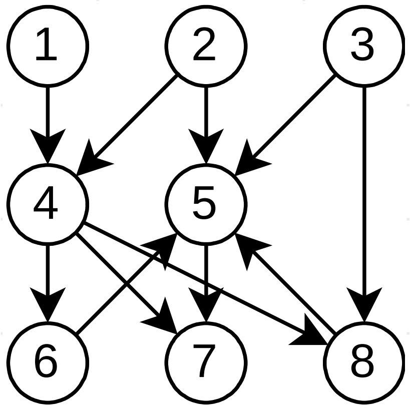
- Phenotype:

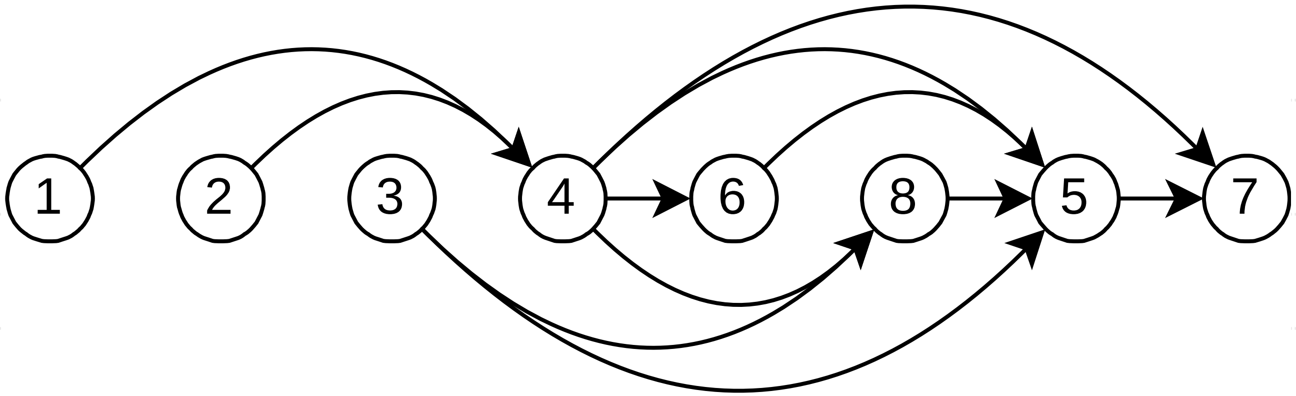


- * **Blue** connections denote data channel with spatial property, **red** connections denote no spatial property

Activation

- Activate in topological order, ensure all signals to a node can be reached before output





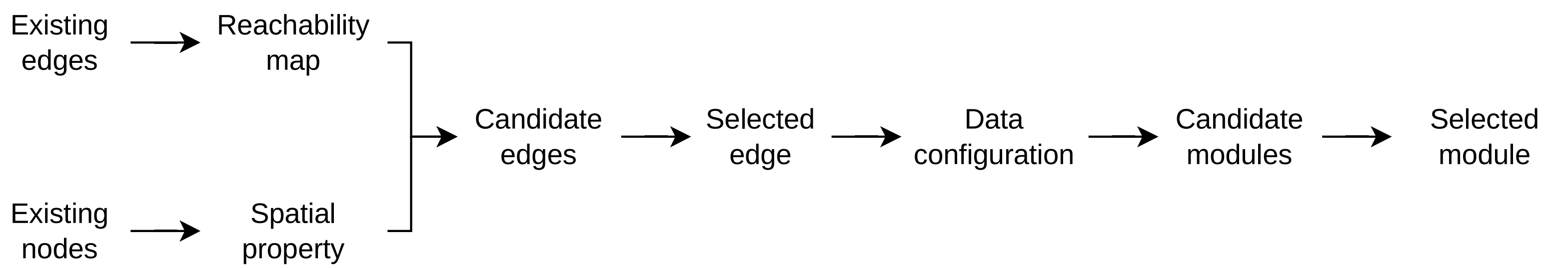
Mutation

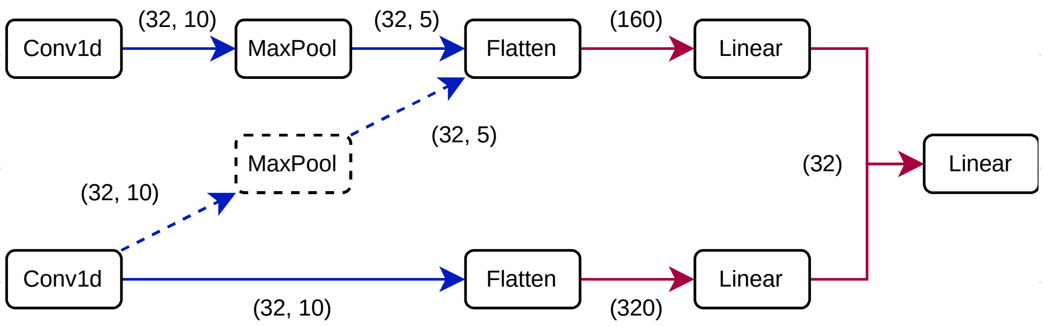
Ensure all possible legal topologies can emerge (theoretically)

- Adding one node
- Adding two nodes
- Adding one edge
- Removing one edge
- Replicating a node

Mutation

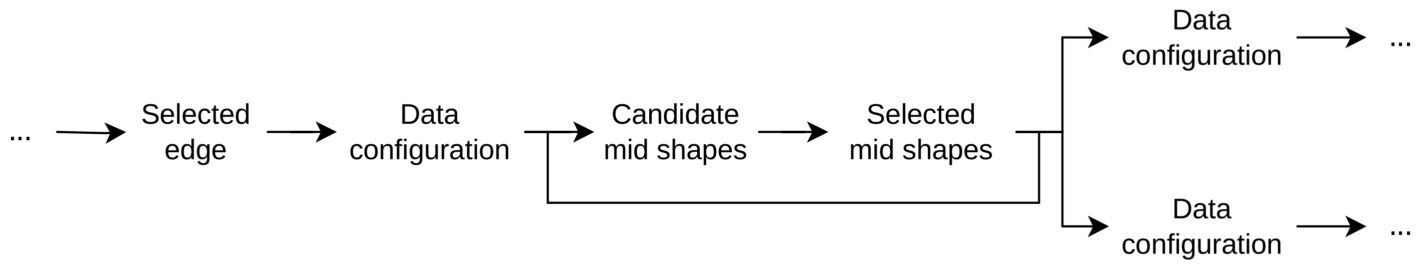
- Adding one node

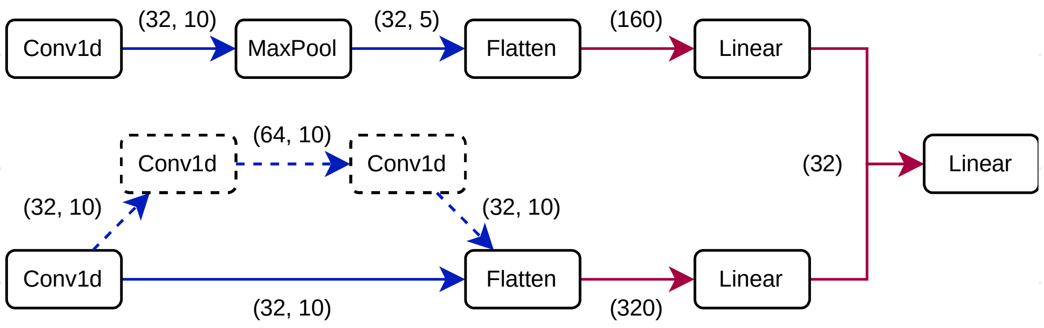




Mutation

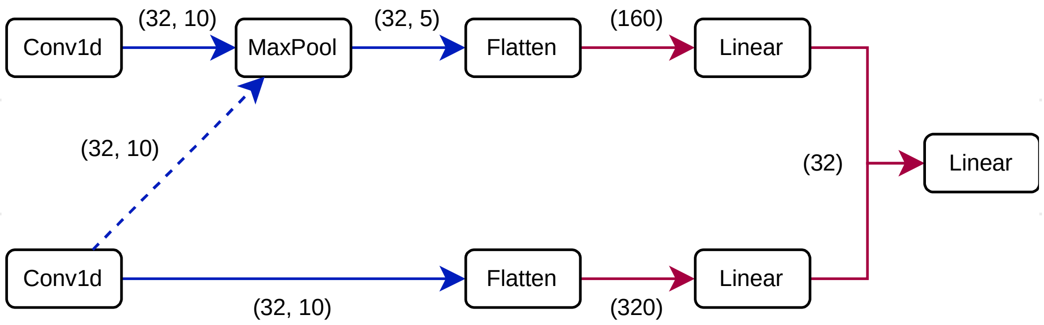
- Adding two nodes





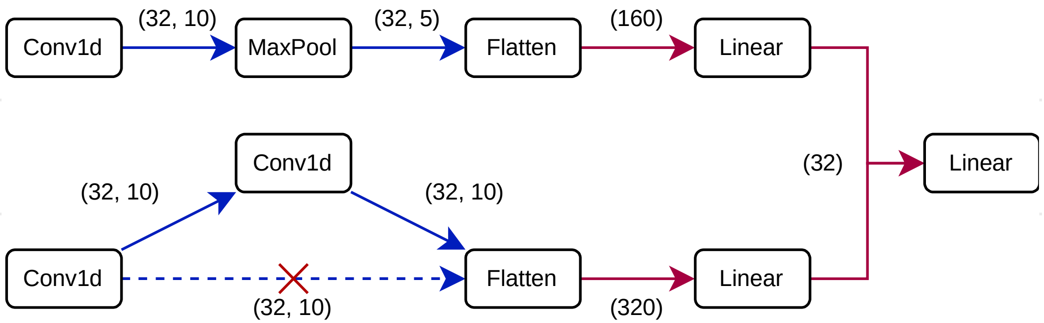
Mutation

- Adding one edge



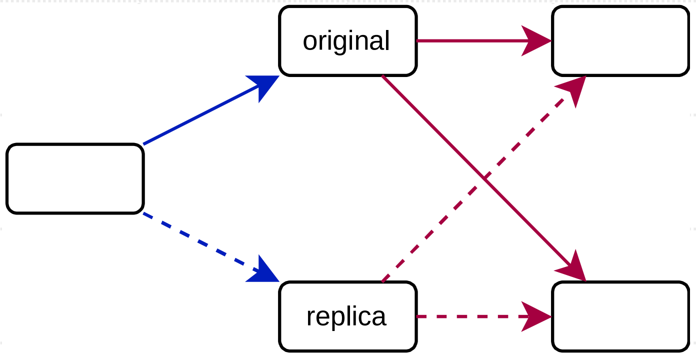
Mutation

- Removing one edge

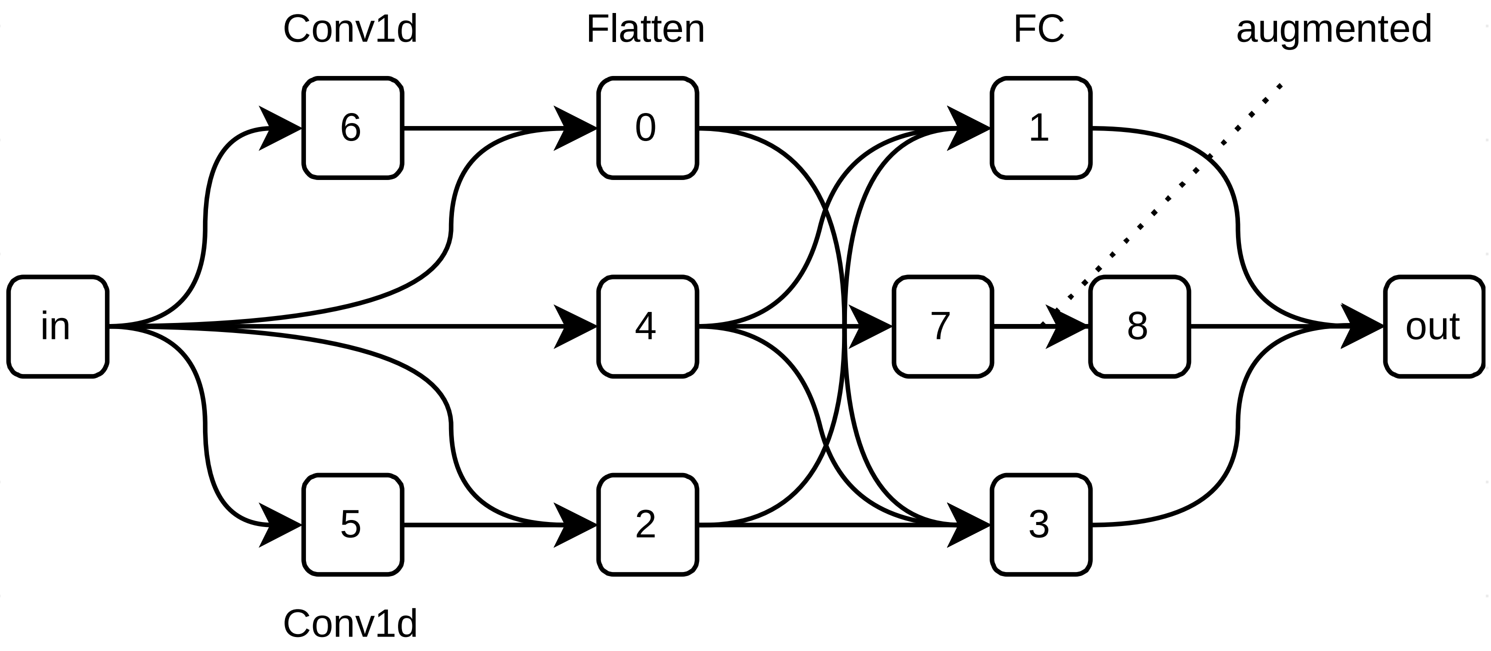


Mutation

- Replicating one node



Example Model



Natural Selection

- Start with minimal structure mapping input data shape to output data shape

- Trained on training dataset
- Evaluated and selected on validation dataset → Overfitting genotypes would be eliminated
- Winners mutate and added to gene pool

- Tested on unseen, future testing dataset (later than training and validation to ensure that the data are unseen)

Trading Strategy

- For simplicity, the model's output is [P(up), P(down)], predicting the stock price's direction in the next day

- If P(direction)>threshold, go all in in that direction, otherwise, hold cash

- Input data: Price, Log Price, RSI, EMA
- Duration: last 32 days

Performance

Trained and evaluated on an individual stock, 5th generation:

Company	Price %	Return %	Outpfm	APY	# of L	# of S	# of H	Hit rate	Sharpe	Sortino	MDD
AAPL	0.74%	56.83%	56.09%	25.57%	303	120	75	56%	1.00	0.86	20%
AXP	22.05%	27.41%	5.37%	13.04%	299	121	78	50%	0.47	0.40	19%
BA	-7.88%	0.16%	8.05%	0.08%	272	144	82	52%	0.16	0.15	24%
CAT	23.57%	29.16%	5.58%	13.82%	272	147	79	50%	0.48	0.40	20%
CSCO	5.03%	21.63%	16.61%	10.42%	275	155	68	54%	0.54	0.52	23%
CVX	95.95%	-21.66%	-117.61%	-11.62%	287	124	87	51%	-0.37	-0.30	40%
DD	-18.28%	50.31%	68.58%	22.90%	269	150	79	45%	1.65	-0.90	23%
DIS	-51.49%	-6.62%	44.87%	-3.40%	288	137	73	50%	-0.01	-0.01	32%
GE	-8.53%	-7.03%	1.50%	-3.62%	262	169	67	49%	0.04	0.04	40%
GS	16.86%	0.76%	-16.10%	0.38%	291	135	72	51%	0.17	0.16	37%
HD	16.45%	31.06%	14.61%	14.67%	301	125	72	49%	0.72	0.64	26%
IBM	14.61%	27.64%	13.03%	13.14%	258	159	81	50%	0.73	0.71	13%
INTC	-48.72%	118.45%	167.17%	48.50%	274	159	65	52%	1.46	1.67	19%
JNJ	10.84%	35.41%	24.57%	16.58%	291	132	75	47%	1.29	1.12	25%
JPM	-2.86%	23.49%	26.35%	11.27%	292	141	65	50%	0.56	0.59	30%
KO	26.71%	-10.74%	-37.46%	-5.59%	287	130	81	50%	-0.27	-0.27	31%

Performance

Trained and evaluated on an individual stock, 5^{th} generation :

Company	Price %	Return %	Outpfm	APY	# of L	# of S	# of H	Hit rate	Sharpe	Sortino	MDD
MCD	23.01%	-4.80%	-27.81%	-2.46%	310	131	57	52%	-0.03	-0.03	24%
MMM	-27.41%	-1.71%	25.70%	-0.87%	289	148	61	53%	0.07	0.06	31%
MRK	36.79%	15.96%	-20.84%	7.78%	295	130	73	50%	0.48	0.49	30%
MSFT	10.27%	-19.62%	-29.89%	-10.46%	298	141	59	52%	-0.35	-0.30	34%
NKE	-20.43%	2.74%	23.17%	1.38%	291	137	70	52%	0.19	0.17	36%
PFE	35.66%	10.28%	-25.38%	5.08%	277	132	89	52%	0.37	0.34	35%
PG	9.95%	-13.95%	-23.90%	-7.32%	284	136	78	47%	-0.60	-0.62	39%
RTX	44.58%	26.07%	-18.51%	12.44%	274	136	88	52%	0.63	0.58	21%
TRV	36.43%	25.95%	-10.47%	12.38%	277	133	88	47%	0.73	0.66	18%
UNH	46.97%	2.08%	-44.89%	1.05%	311	116	71	48%	0.17	0.17	27%
VZ	-31.43%	-14.54%	16.89%	-7.64%	283	137	78	50%	-0.42	-0.37	25%
WBA	-21.68%	-18.05%	3.62%	-9.58%	277	151	70	51%	-0.27	-0.23	33%
WMT	-3.73%	95.88%	99.62%	40.53%	297	118	83	54%	1.70	2.00	17%
XOM	135.48%	112.63%	-22.85%	46.48%	264	150	84	51%	1.86	1.75	11%
Total	12.32%	19.84%	7.52%	8.50%	284.93	138.13	74.93	50.58%	0.44	0.35	27.78%

Performance

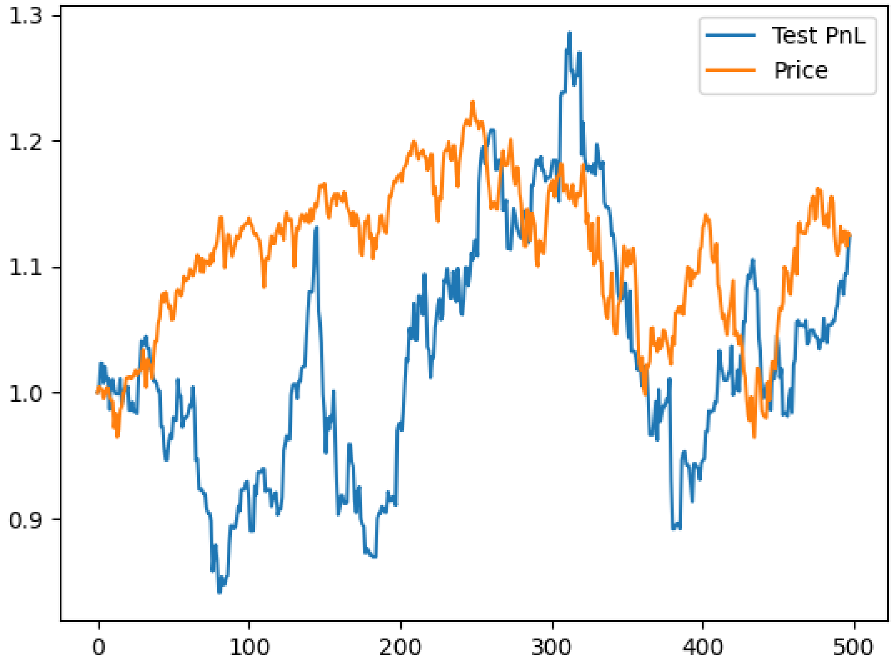
Trained and evaluated on an individual stock:

- Slightly outperform strong hold strategy
- Has a huge MDD
- Accuracy ~ 50%
- Low Sharpe and Sortino

Performance

Trained and evaluated on all stocks, build a portfolio

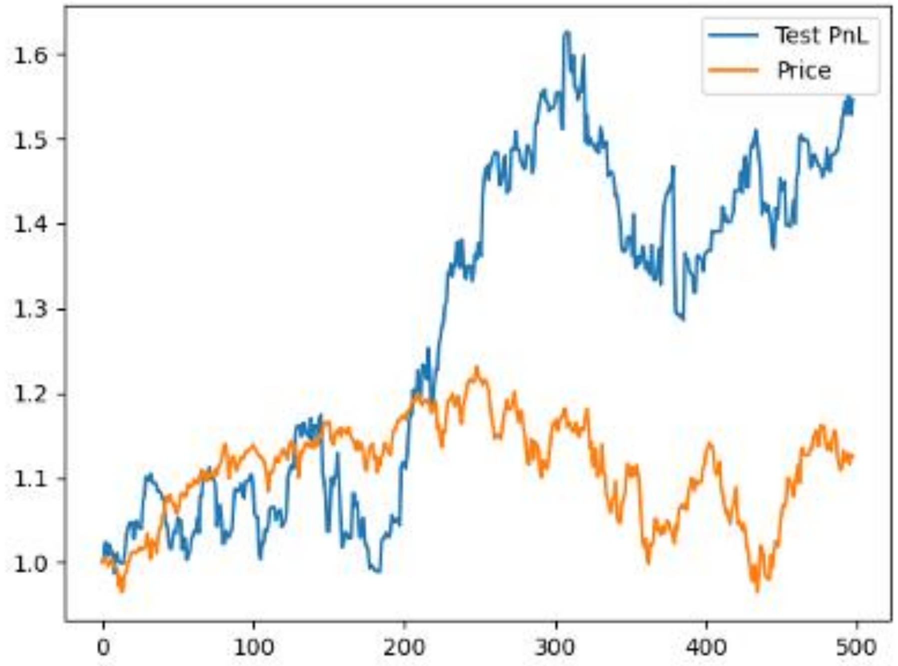
- Generation 0
- Population: 1



Performance

Trained and evaluated on all stocks, build a portfolio

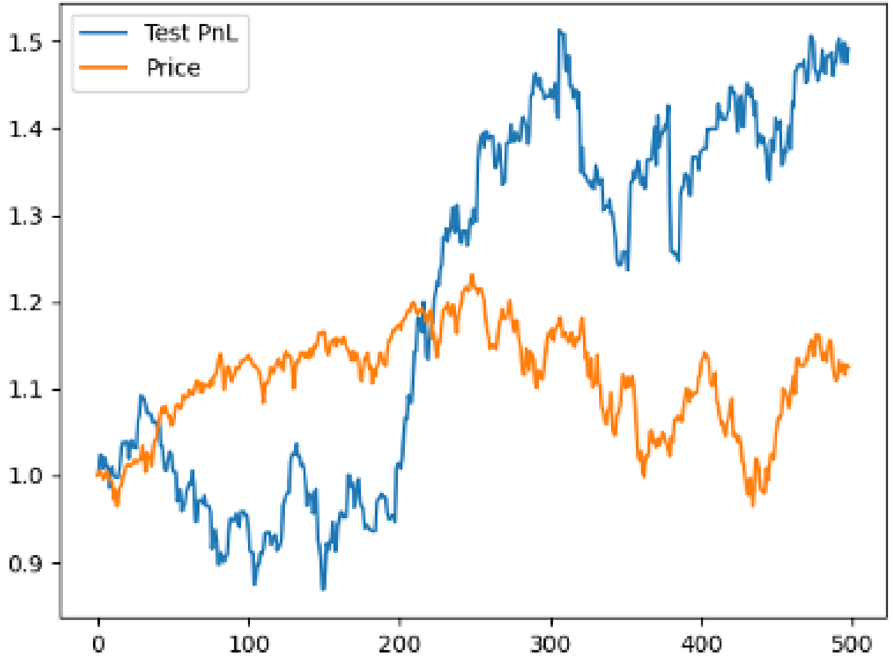
- Generation 1
- Population: 4



Performance

Trained and evaluated on all stocks, build a portfolio

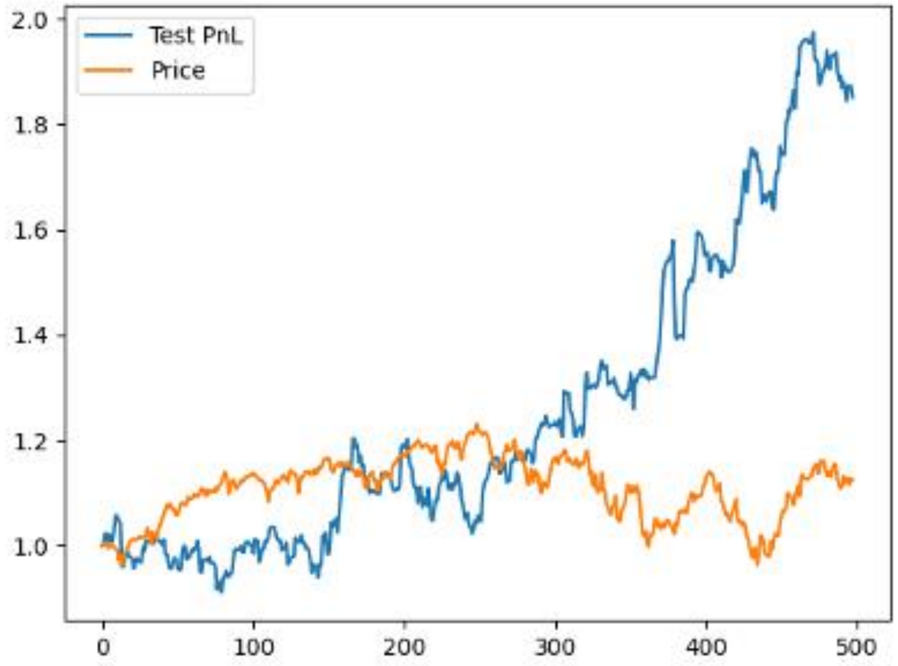
- Generation 2
- Population: 16



Performance

Trained and evaluated on all stocks, build a portfolio

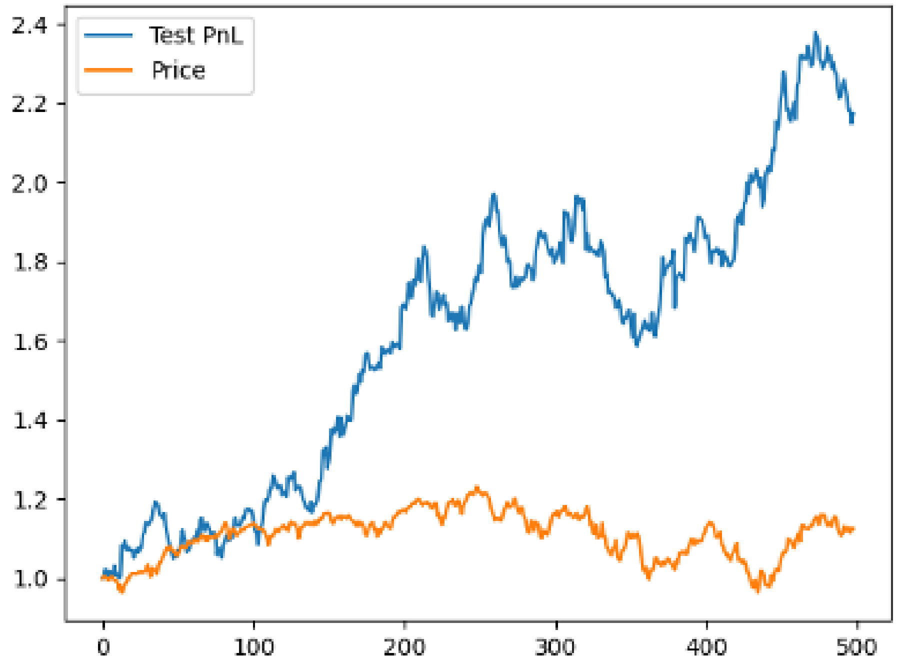
- Generation 3
- Population: 32



Performance

Trained and evaluated on all stocks, build a portfolio

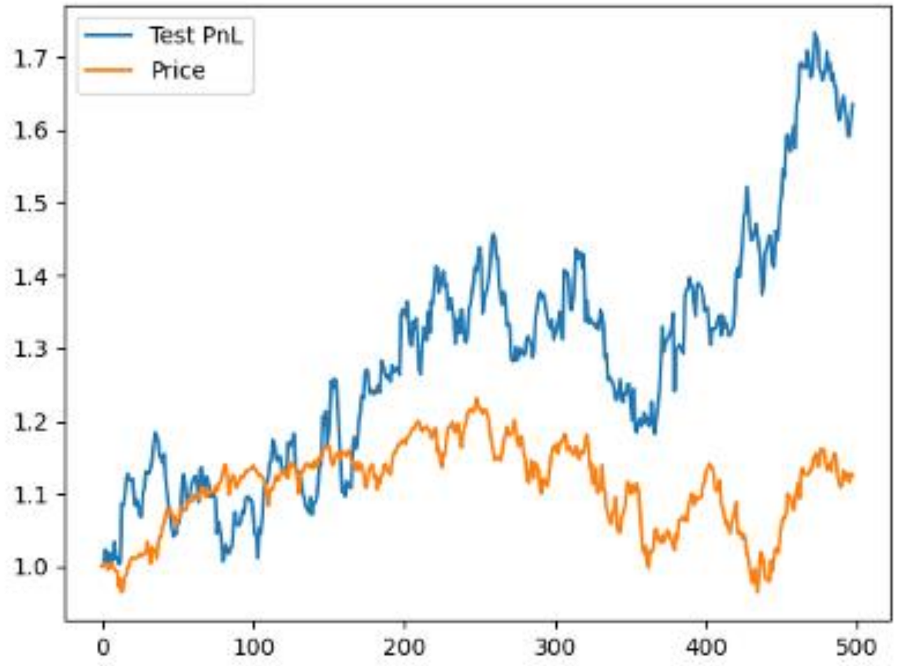
- Generation 4
- Population: 32



Performance

Trained and evaluated on all stocks, build a portfolio

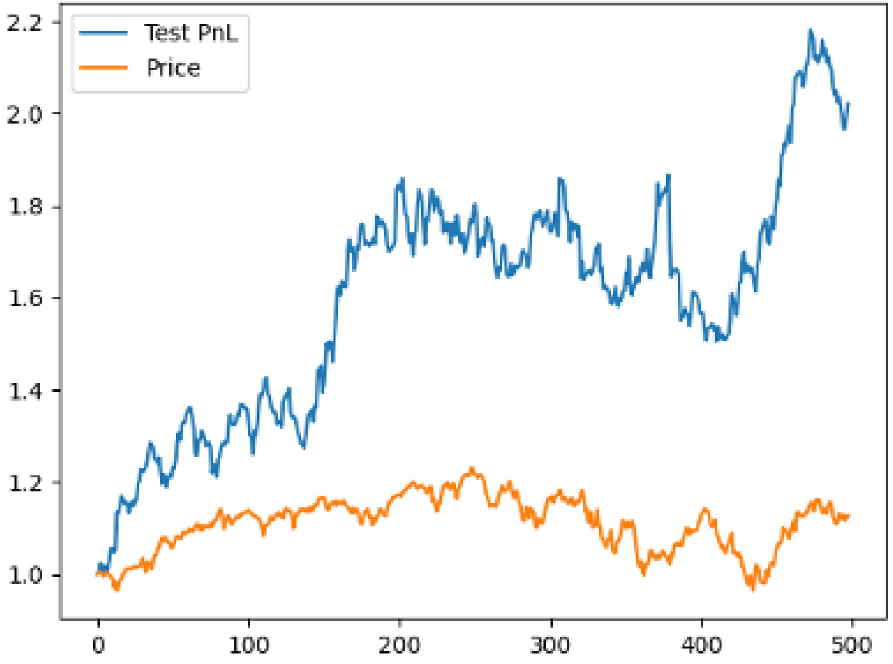
- Generation 5
- Population: 32



Performance

Trained and evaluated on all stocks, build a portfolio

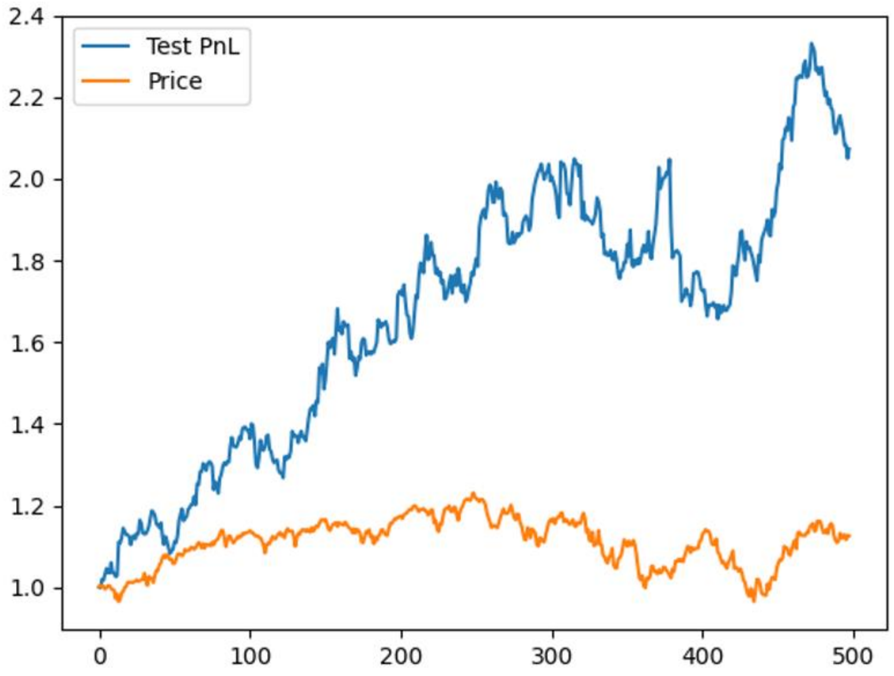
- Generation 6
- Population: 32



Performance

Trained and evaluated on all stocks, build a portfolio

- Generation 7
- Population: 32



Future Works to Add in Final Paper

- Ablation study
- Increase the population size
- Increase the number of generations
- Correlation between model complexity and performance
- Add more rules to natural selection, e.g., takes Sortino ratio into account
- Testing different thresholds