Pioneering Intelligence

Initiative of Flagship Pioneering

Flow-of-Options: Diversified and Improved LLM Reasoning by Thinking Through Options

Lakshmi Nair, Ian Trase, and J. Mark Kim

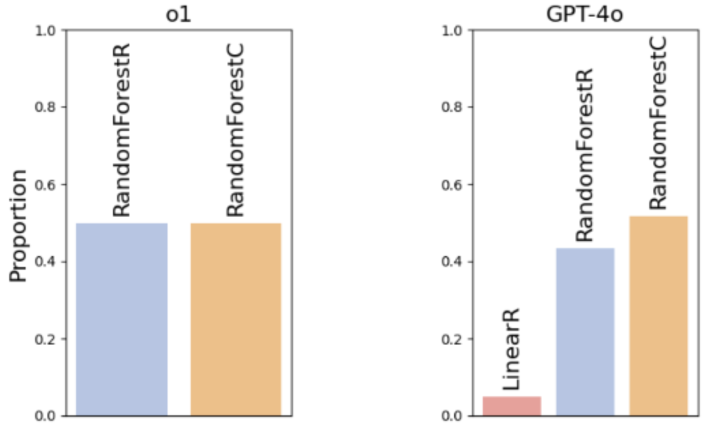
ICML
International Conference On Machine Learning

Motivation – LLM reasoning is not diverse – LLMs have biases

We focus on this issue in the context of **automated design of ML pipelines (AutoML)** – LLMs show pre-training biases for certain ML solutions compared to others:

User provides task

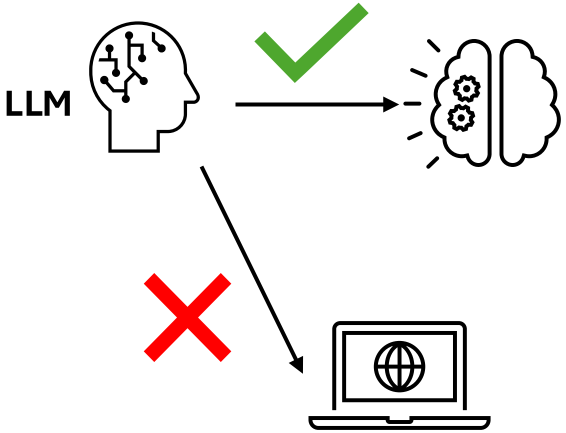
“I would like to perform a regression/classification task, where given a drug SMILES string, I’d like to predict…”



Not great for all ML problems
We need “<u>diversity</u>”

LLM almost always prefer RandomForest in its proposed solutions

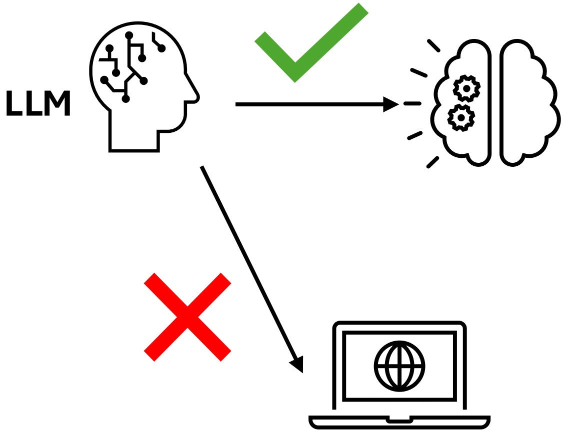
LLMs have a lot of useful information; but are bad at using it



LLMs have knowledge on different ML models

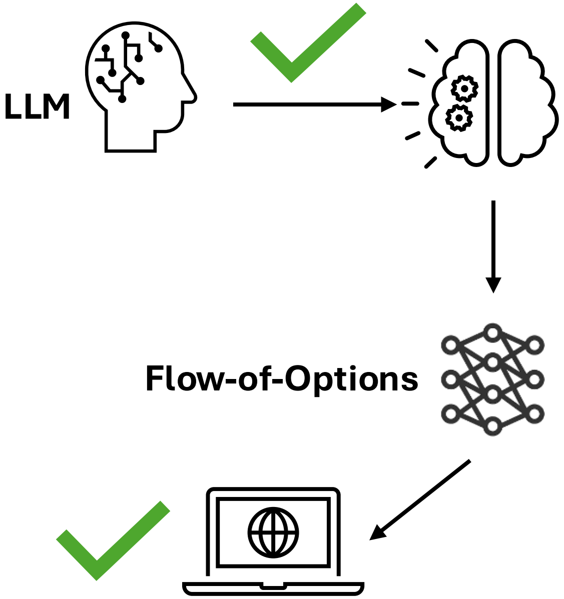
However, they always *blindly* choose one class of models when writing actual code, without considering the different choices and their implications

LLMs have a lot of useful information; but are bad at using it



LLMs have knowledge on different ML models

However, they always *blindly* choose one class of models when writing actual code, without considering the different choices and their implications

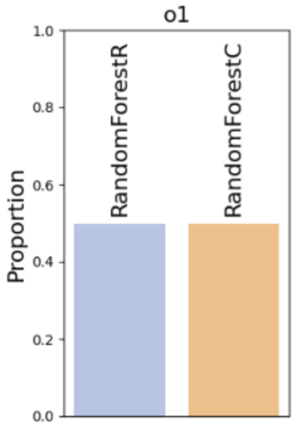


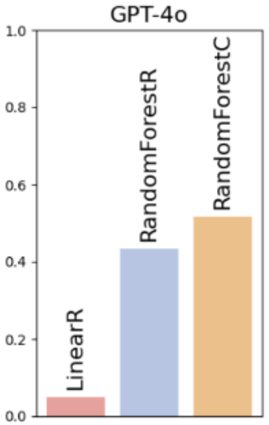
Use the LLM's *thinking* but let Flow-of-Options do the *implementing*

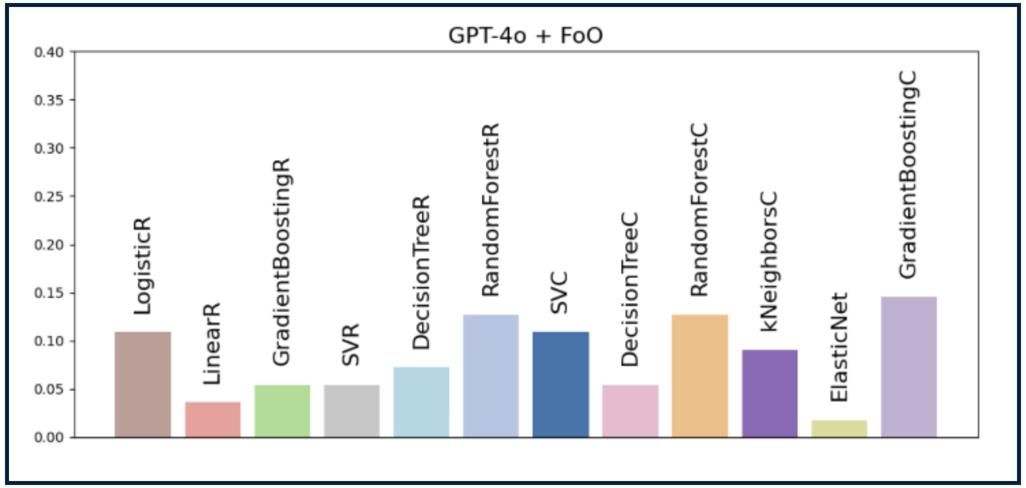
Flow-of-Options

Flow-of-Options
A "fully-connected" network data structure that captures different "choices" in the ML pipeline, that can then be systematically explored to identify the best ones

User provides task
"I would like to perform a regression/classification task, where given a drug SMILES string, I'd like to predict…







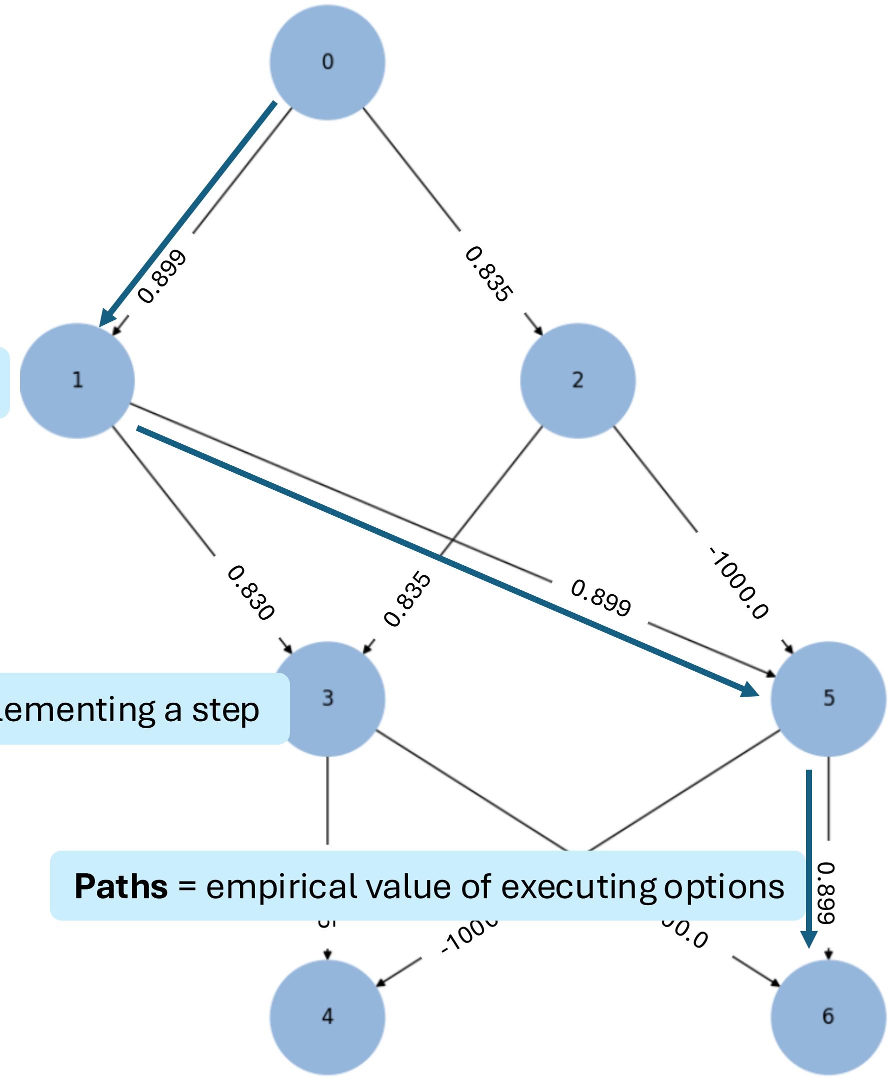
Flow-of-Options

Looks like a fully-connected neural network

Levels = task plan step

Nodes = option for implementing a step

Paths = empirical value of executing options



0:

1: Convert the SMILES strings into Morgan fingerprints (circular fingerprints) using RDKit, specifying a radius and bit length, and use these fingerprints as input features for the model.

2: Use Morgan fingerprints (also known as circular fingerprints) generated by RDKit to convert SMILES strings into binary vectors. These vectors can be used as input features for the model, capturing structural information about the molecules.

3: Use RandomForestClassifier from sklearn with default parameters to handle the binary classification task. This model is robust to overfitting and can handle the complexity of molecular descriptor features effectively.

4: Train the RandomForestClassifier using the preprocessed training data and corresponding labels by employing a stratified k-fold cross-validation approach on the training set to ensure robust performance and mitigate any potential overfitting.

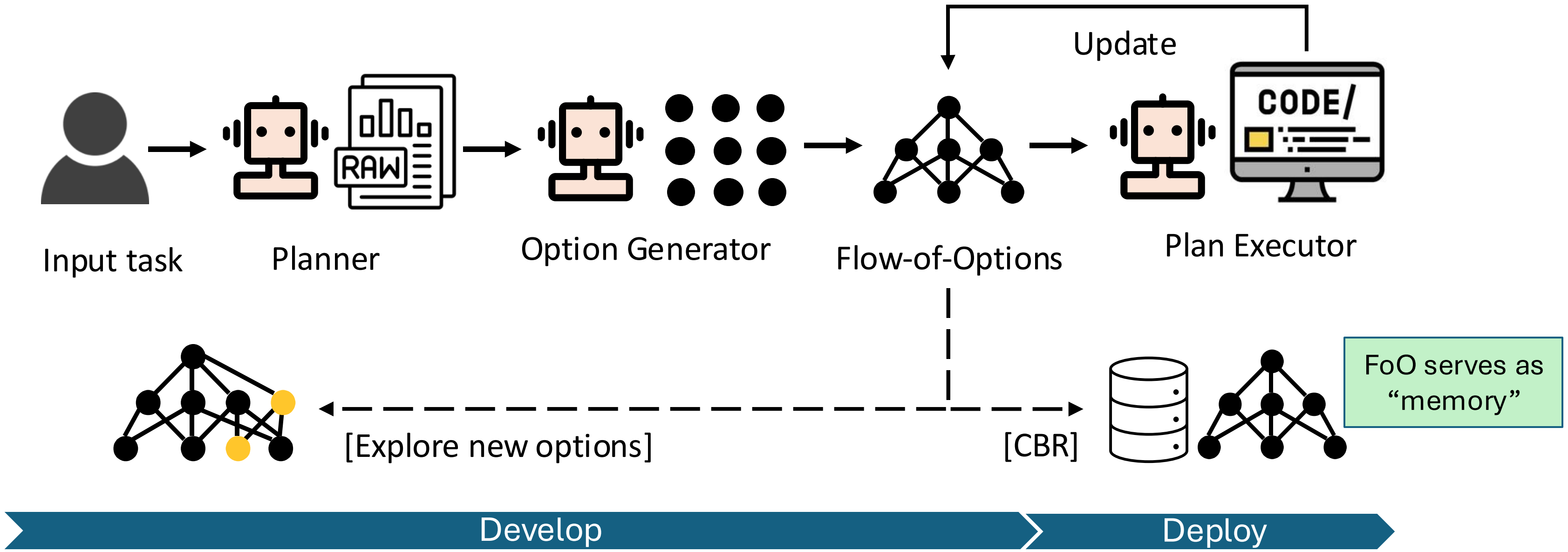
5: Use LogisticRegression from sklearn with L2 regularization and a maximum of 100 iterations to handle the binary classification task. This model is simple and effective for linear decision boundaries.

6: Train the LogisticRegression model using the preprocessed training data and corresponding labels with the addition of early stopping based on validation loss to prevent overfitting and ensure the model generalizes well to unseen data.

Agentic System for ML using Flow-of-Options

We combine Flow-of-Options with Case-based Reasoning (CBR)

Benefits of FoO + CBR:
- Quick and efficient deployment (low cost and low time)
- **Condensed memory** of prior choices that can be improved upon



Results – FoO outperforms benchmarks

Flow-of-Options outperforms state-of-the-art AutoML benchmarks

Typical Data Science Tasks

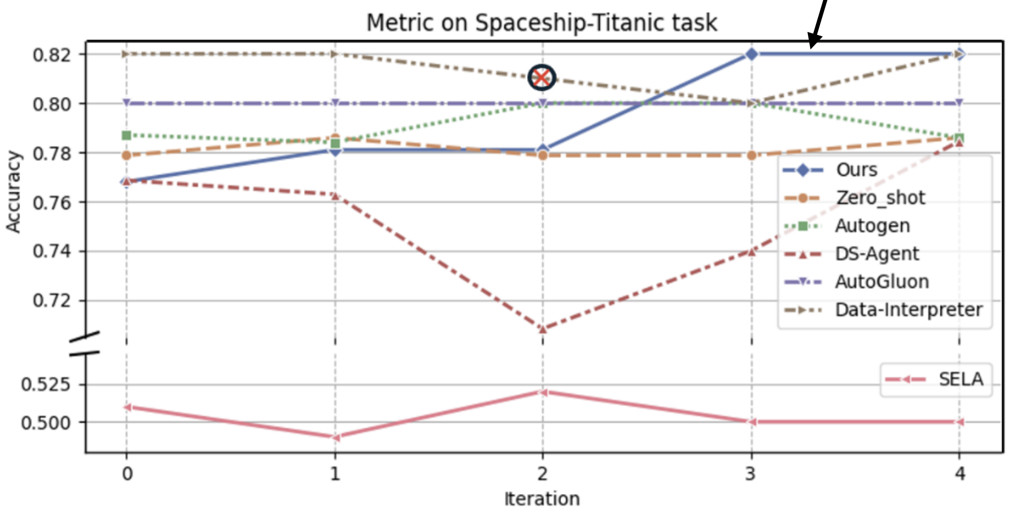
	Development							Deployment									Avg.
	WB (↓)	MC (↓)	ES (↑)	EC (↑)	AR (↓)	ST (↑)	ILI (↓)	SS (↑)	MH (↓)	W (↓)	SD (↑)	J (↓)	CA (↓)	CS (↓)	HB (↑)	WR (↑)	Rank (↓)
DS-Agent	304	0.30	0.40	0.27	4.47	0.78	6.49	**0.99**	0.34	**0.34**	0.80	**0.67**	0.73	11.7	0.68	0.65	3.69
AutoGluon	322	0.28	0.61	–	–	0.80	–	0.89	0.54	–	0.69	–	1.36	11.4	–	–	4.67
SELA	321	0.29	0.71	–	1.19	0.51	–	0.72	0.32	–	0.85	0.81	1.39	11.8	–	0.75	4.17
DI	314	0.30	**0.98**	**0.43**	**1.11**	**0.82**	**1.05**	0.88	**0.06**	X	0.82	0.98	**0.40**	9.88	**0.76**	0.75	2.33
Autogen	309	0.30	0.67	0.37	1.67	0.80	2.86	0.90	0.40	0.52	0.85	0.69	1.38	10.3	0.72	**0.83**	3.19
Zero-shot	263	0.26	0.80	0.35	1.91	0.78	5.19	0.83	0.50	0.38	0.81	0.79	1.16	10.0	0.72	**0.83**	3.19
Ours	**182**	**0.18**	**0.98**	**0.43**	1.59	**0.82**	1.53	**0.99**	0.29	0.36	**0.98**	**0.67**	0.73	**9.18**	**0.76**	0.80	**1.44**

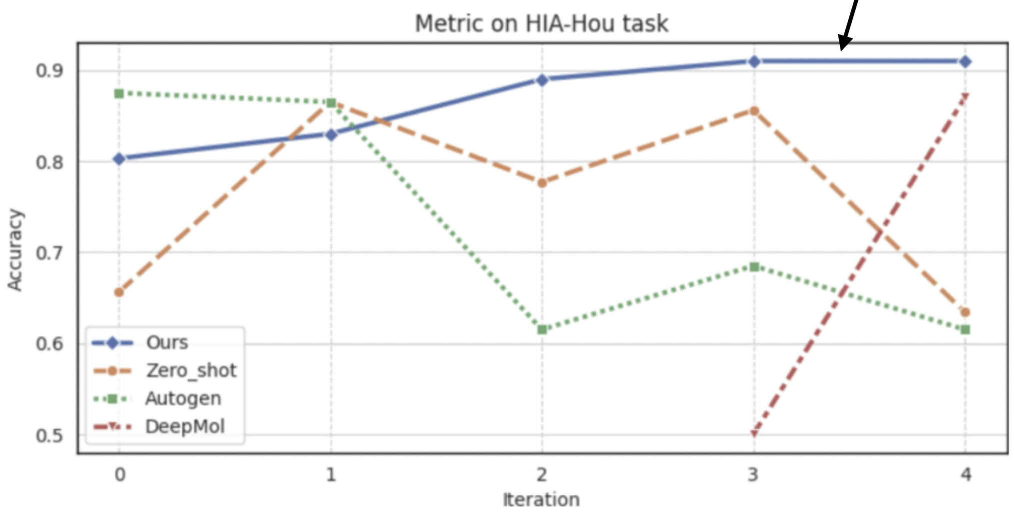
Therapeutic Data Commons (ADME-Tox prediction)

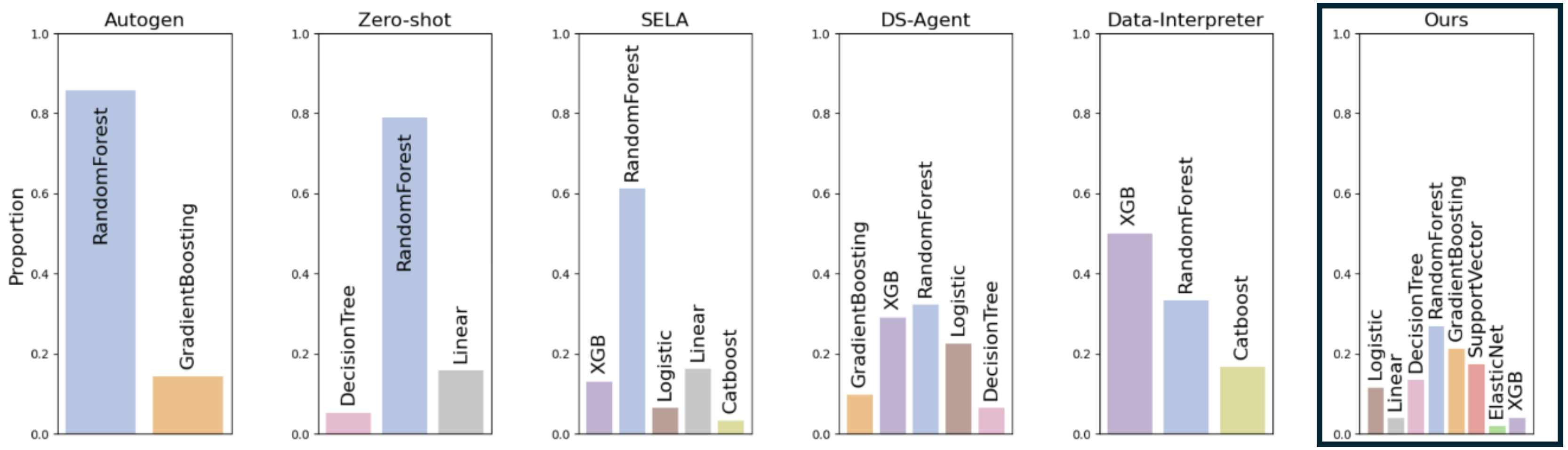
	Development								Deployment									Avg.
	CW (↓)	HH (↑)	BI (↑)	PG (↑)	LI (↓)	BB (↑)	PP (↓)	VD (↑)	C2 (↑)	C3 (↑)	C2S (↑)	C3S (↑)	HO (↑)	CH (↑)	A (↑)	hE (↑)	DI (↑)	Rank (↓)
DeepMol	0.35	0.87	0.50	0.82	**0.69**	0.72	**8.34**	0.33	0.20	0.64	0.38	**0.63**	**0.49**	0.14	0.65	**0.76**	0.79	2.35
Autogen	0.42	0.77	0.50	0.86	0.77	0.73	10.3	0.43	0.36	0.63	0.42	0.59	0.32	0.28	**0.76**	0.64	0.72	2.71
Zero-shot	0.44	0.81	0.56	0.81	0.76	0.77	11.2	0.29	0.34	0.64	0.39	0.59	0.28	**0.32**	0.71	0.69	0.72	2.82
Ours	**0.34**	**0.91**	**0.58**	**0.89**	0.75	**0.78**	9.51	**0.52**	**0.57**	**0.80**	**0.62**	0.62	0.26	**0.32**	**0.76**	0.68	**0.84**	**1.47**

FoO → improvement over time + diversity

Flow-of-Options shows **capacity to improve** over time by building on past options (CBR)







Explores **a broader range of options**

FoO can generalize beyond tabular ML

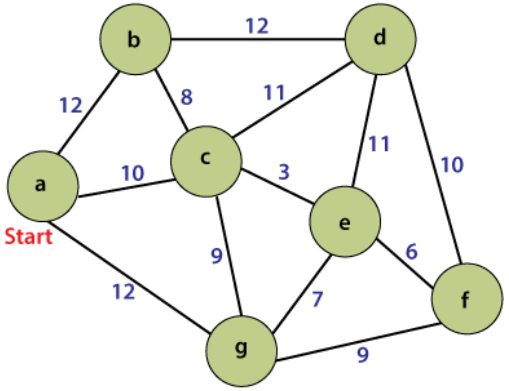
Sequence-to-sequence translation

Translation from English to French

English input: ”The fool doth think he is wise, but the wise man knows himself to be a fool.”

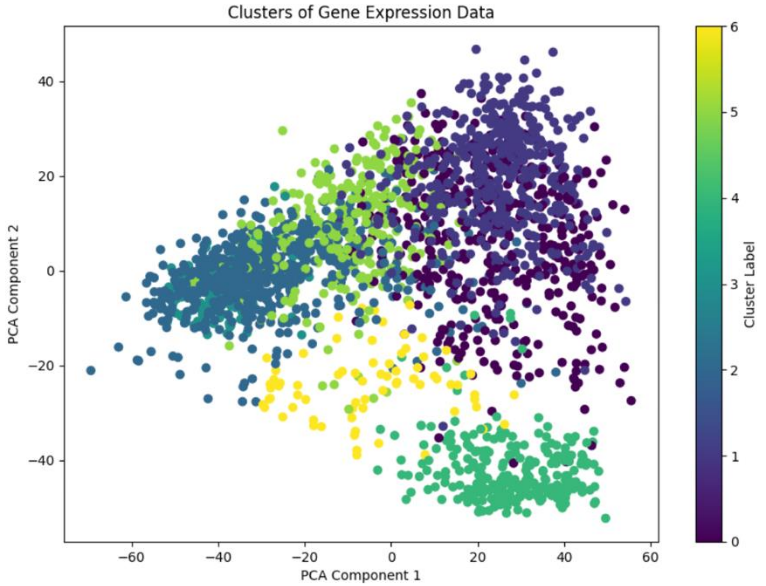
French output: "Le bête pense qu'il est sage, mais le sage se sait être un bête."

Traveling Salesman Problem



Clustering

Clustering of Gene Expression Data



Reinforcement Learning

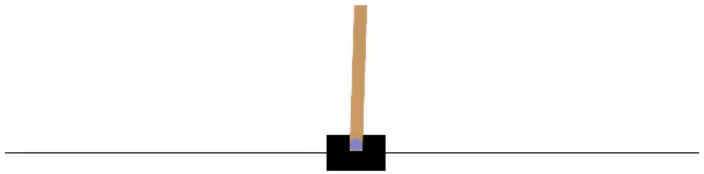
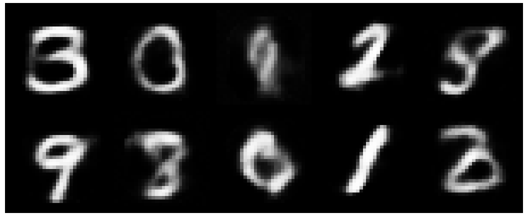


Image Generation



Math Task

The equation $x^2 + 2x = i$ has two complex solutions. Determine the product of their real parts

Pioneering Intelligence

Initiative of Flagship Pioneering

Thank you!

GitHub: https://github.com/flagshippioneering/Flow-of-Options