Effect of bio-plausible learning rules on MLP's weights

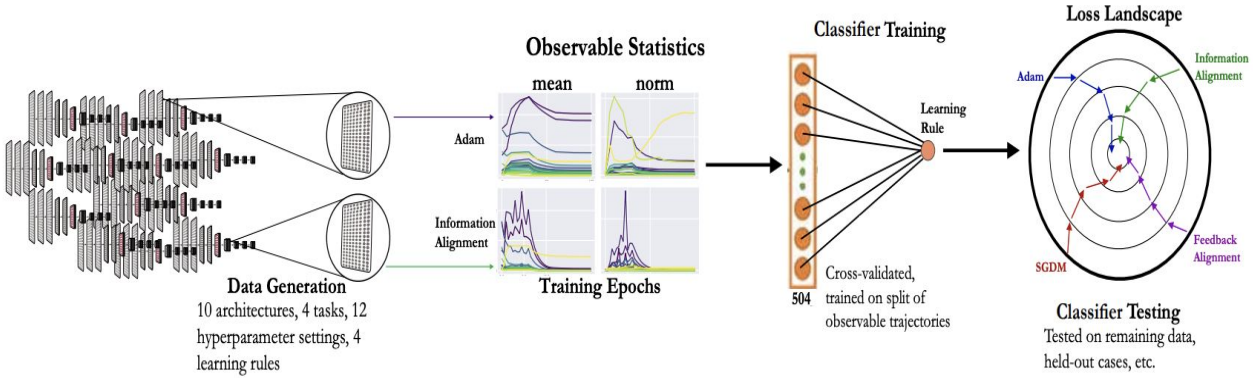
Microlearners
Ali Mohammadnejad, Mina Rezaei, Tamaz Gadaev, Fatemeh Shahbodaghy, Safa Mohammadi

Regular TA: Alish Dipani
Project TA: Aakash Agrawal

The Why: Research questions

- How does bio-plausible learning rules affect resulting weights?

Experiment design



Experiment is based on [1]

[1] Identifying Learning Rules From Neural Network Observables

Experiment design

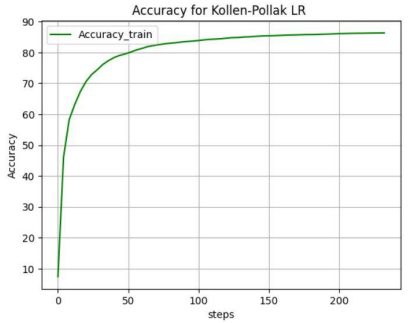
Same parameters for all the networks:

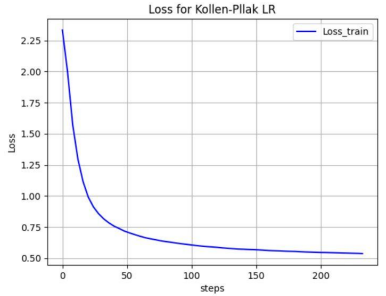
- Hidden size = 100
- Batch size = 128
- Number of epochs = 3
- Learning rate = 0.01 (where possible)
- Initialization variance = 0.1
- No bias in Linear layers

- ❖ Frobenius norm, mean, variance, skew, kurtosis, median, and third quartile (Raw, Abs, Square)
- ❖ Diversity of architectures and hyperparameters in comparison to original paper was limited due to constrained resources

Implemented Learning rules

1. Backpropagation (particularly, SGD algorithm)
2. No training
3. Kollen-Polack
4. Feedback Alignment
5. Hebbian Rule
6. Contrastive Hebbian Rule





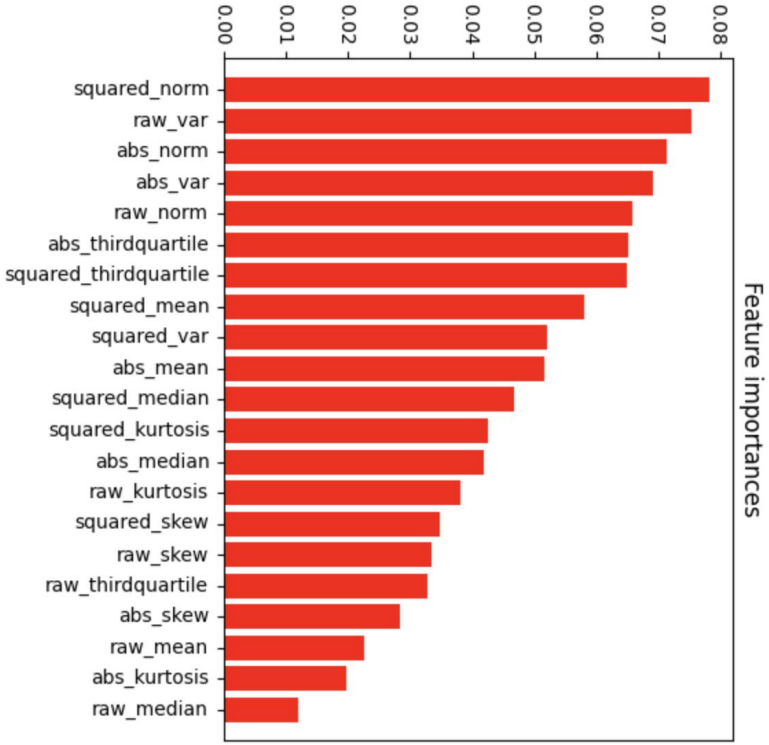
Classification results

Overall Accuracy of Learning Rules classification is 0.75

Classification results

Performance of Random Forest on different learning rules by class

	Precision	Recall	F1	Support
BackProp	67	86	75	7
No training	25	20	22	5
ContrastiveHebb	100	100	100	7
FeedbackAlign	25	20	22	5
Hebbian	100	100	100	5
Kollen-Polak	100	100	100	7



Next steps and Discussion

- Our goal was to understand how different learning rules affect the weight distribution and their classification potential.
- Our approach doesn't fully answer the question, because
  - Narrow scenario: limited number of architectures, training scenarios and hyperparameters, only shallow MLP network
  - Limited number of network features: no activations and gradients
- This research could be extended in these directions:
  - Analyze activation patterns
  - Generalization performance
  - Investigate network dynamics in time

Thanks for your Attention

Questions?

Our Pod: Appreciative Nemesia