A Meta-cognitive Recurrent Fuzzy Inference System with Memory Neurons (McRFIS-MN) and its Fast Learning Algorithm for Time Series Forecasting

2018 Symposium Series on Computational Intelligence
IEEE SSCI

Subhrajit Samanta **(PhD student, IGS NTU)**
Shubhangi Ghosh **(B.Tech student, IIT Madras)**
Suresh Sundaram **(Professor, IISC Bengaluru)**

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Background and Motivation

Time Series Modeling [11]:

Set of data points which possess uncertainty and a sequential nature

Features of McFIS-MN to model time series :

- Meta-Cognition **(Mc)**
  To control the learning process;
  by deciding what-to-learn, when-to-learn and how-to-learn
- Fuzzy Inference System **(FIS)**
  To handle uncertainty
- Memory Neurons **(MN)**
  To handle system dynamics by introducing recurrence
- Projection-Based Learning **(PBL)**
  To ensure a fast one shot learning

[11] Box Jenkins Book (2015)

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Literature Review

Fuzzy Inference Systems (FIS):

1. Fixed Structure FIS [1]:
   - Handles uncertainty well with human-like linguistic behavior
   - Not efficient in handling non-stationary data
   - Not efficient in handling sequential data

2. Self-Adaptive FIS [2-4]:
   - Handles uncertainty well with human-like linguistic behavior
   - Capable of handling non-stationarity using Meta-cognitive evolving stucture
   - Not efficient in handling sequential stream of data with BP

Learning Methods:

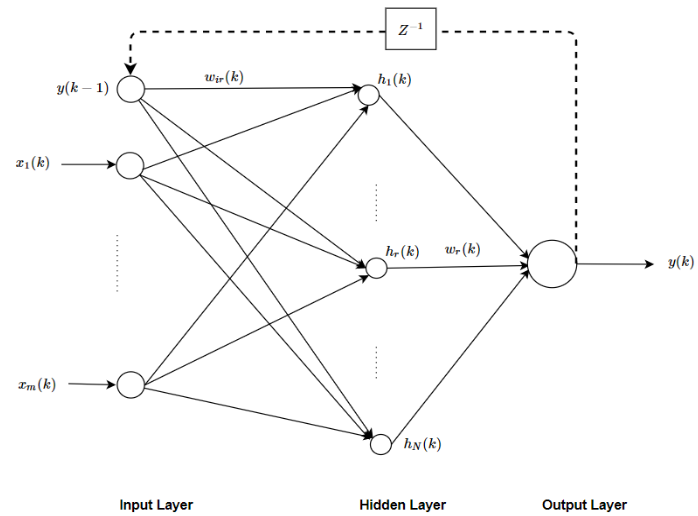
1. Backpropagation (BP) Through Time [5]:
   - Slower, needs multiple epochs hence not suitable for sequential learning
2. Projection-Based Learning [6]:
   - Faster one-shot learning method suitable for sequential learning

[1] Takagai and Sugeno (1993)) [2] SAFIS (2006) [3] McFIS (2012) [4] PANFIS (2014) [5] BPTT Proceedings of the IEEE (1990) [6] S. Suresh et. al. (2013)

Literature Review

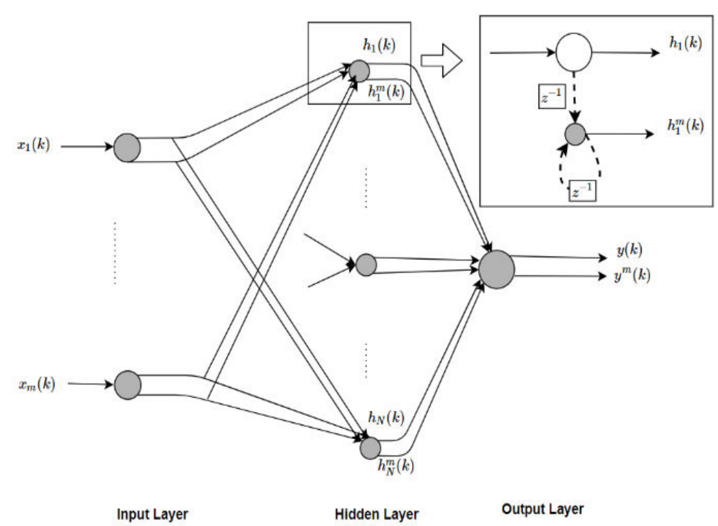
Recurrent Structures:

1. Regular Feedback Structures [7]:
   - Can handle sequential behavior
   - Can't handle unobservable dynamic systems
   - System order needs to be known
   - Stability can be a critical issue



1. Memory Neuron Network [8]:
   - Can handle system dynamics well
   - System order doesn't need to be known
   - Stability isn't an issue

Memory Neuron

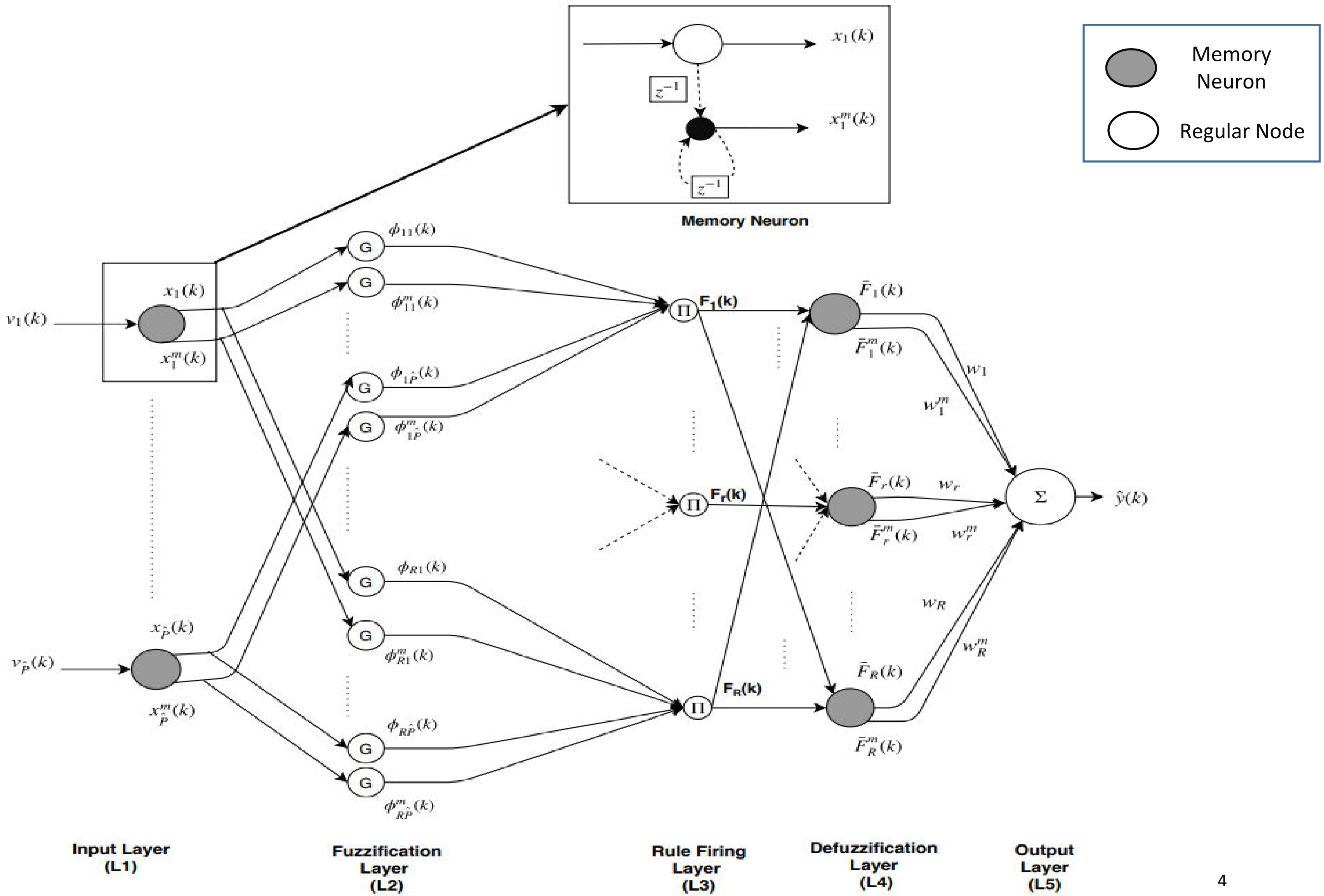


[7] Levin et al.(1996) IEEE TNN *[8] P. S. Sastry et al. (1994) IEEE Transaction on Neural Networks*

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Proposed McRFIS-MN structure



Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Meta-cognitive Structure Learning

The prediction error - e(k) is used to decide one among the following strategies for handling the k^th training sample

1. **Sample deletion:** If $|e(k)| < E_d$ then delete k^{th} sample

2. **Sample Learning:** $|e(k)| > E_d$

Rule Addition	Weight Update	Rule Pruning
• If error is higher than Add Threshold • Spherical Potential is lower than a threshold. If $\|e(k)\| > E_a$ and $\Psi < E_S$ then add a new rule • A lower potential indicates higher novelty **hence add a new rule** • If not novel go for weight update	• When the sample isn't very novel, • weights are tuned using **PBL** • Discussed in next slide	• If the contribution of a rule is low for more than a certain number of samples the rule is pruned. • Ensures a compact network size

3. **Sample Reserve:** If none of the above conditions are satisfied, sample is deferred for learning at a later stage

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Projection-Based Parameter Learning

- **Cost function = Sum of squared error for all the training samples**

$$J(\mathbf{w}) = \frac{1}{2}\Sigma_{k=1}^{S} e^2(k)$$

- **Goal is to minimize the cost function,**

$$\mathbf{w}^{\star} = argmin J(\mathbf{w})$$

- **Equating the partial derivative of J(w) with respect to w to 0 and rearranging we have,**

$$A\mathbf{w} = B$$

- **Solution provides the optimal weights as,**

$$\mathbf{w}^{\star} = A^{-1}.B$$

Where,

$$a_{rr^*} = \Sigma_{k=1}^{S} \hat{F}_r(k).\hat{F}_{r^*}(k)$$

$$b_r = \Sigma_{k=1}^{S} \hat{F}_r(k).y(k), \quad r, r^* = 1, 2, \cdots, 2R$$

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Experimental Results

- Standard nonlinear dynamic system identification problems
  - Nonlinear dynamic SID problem 1 [2,9]

Synthetic, nonlinear, dynamic

- Benchmark time series forecasting problems
  - Mackey Glass chaotic time series problem [10]
  - Box Jenkins CO2 emission prediction problem [11]
  - Sunspot number prediction problem [11]

Nonlinear, dynamic with uncertainty

[9] Narendra et al. (1990) IEEE TNN [2] H.J.Rong et al. (2006) Elsevier [10] R.S.Crowder(1991) Elsevier [11] Box Jenkins Book (2015)

Experimental Results

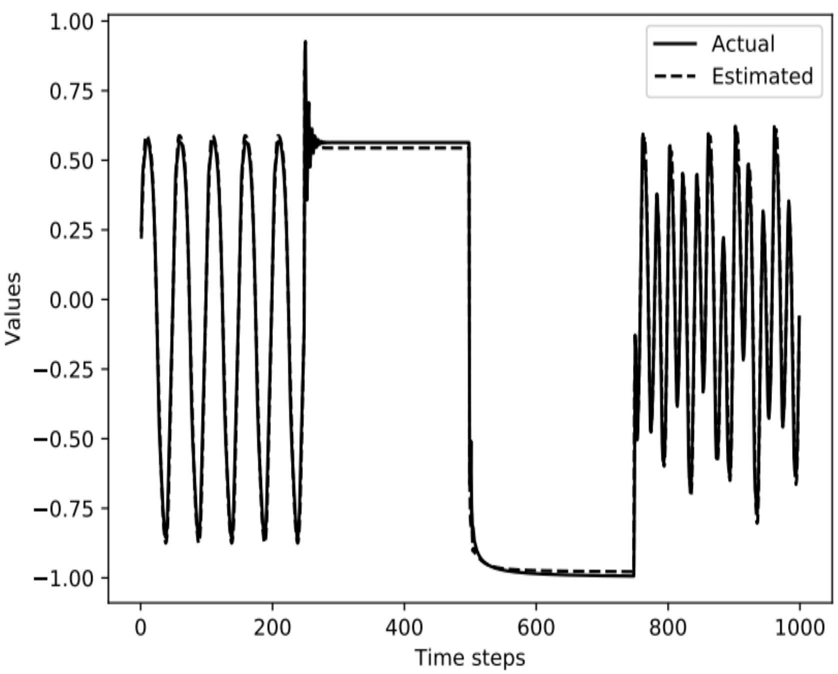
Nonlinear dynamic system identification problems

- **900 samples for training, 1000 samples for testing**

Performance Comparison

Problem	Network	#Rules	Testing RMSE	CPU time(s)
SID1	eTS	49	0.021	3
	SimpleTS	22	0.030	5
	SAFIS	17	0.022	4
	McFIS	10	0.030	7
	McRFIS-MN	**14**	**0.040**	**0.46**

Actual vs Estimated Plot



Experimental Results

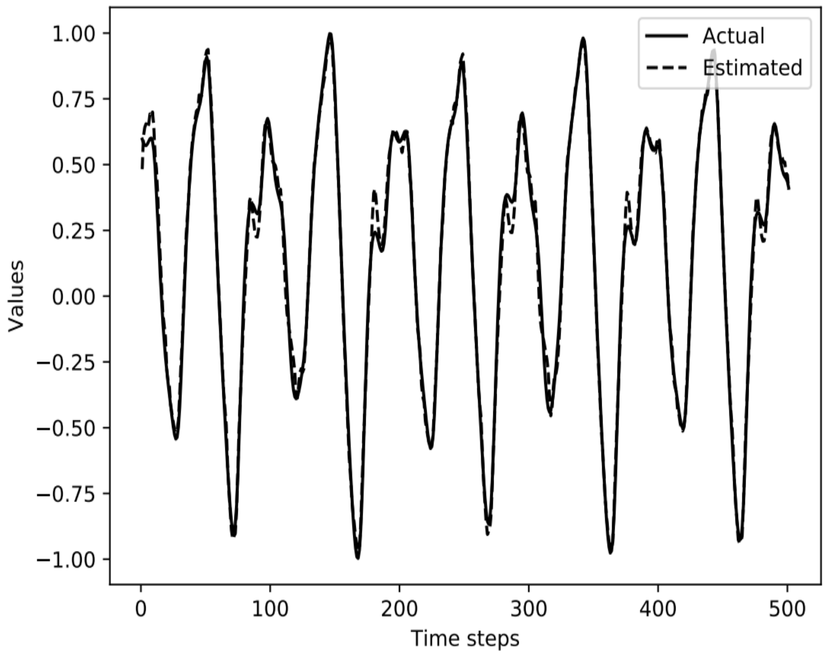
Mackey Glass chaotic time series problem

- **3000 samples for training, 500 for testing**

Performance Comparison

Problem	Network	#Rules	Test NDEI	CPU time(S)
MG	eTS	9	0.380	0.3
	SAFIS	6	0.376	0.5
	SimpleTS	11	0.394	0.4
	McFIS	10	0.100	0.9
	McRFIS-MN	**14**	**0.110**	**0.3**

Actual vs Estimated Plot



Experimental Results

Box Jenkins and Sunspot time series problem

Performance Comparison

Problem	Network	#Rules	Test RMSE	CPU time(S)
BJ	eTS	9	0.049	0.4
	SAFIS	5	0.071	0.6
	SimpleTS	5	0.049	3
	McFIS	12	0.036	0.2
	McRFIS-MN	**5**	**0.033**	**0.04**
Sunspot			**Test RMSE**	
	eTS	23	0.047	3.5
	SAFIS	21	0.100	4.4
	SimpleTS	20	0.050	3.2
	McFIS	12	0.060	4.2
	McRFIS-MN	**5**	**0.044**	**0.15**

Outline

- Background and Motivation
- Literature Review
- Methodology
  - Proposed McRFIS-MN structure
  - Meta-cognitive Structure Learning
  - Projection-Based Parameter Learning
- Experimental Results
- Conclusion and Future Directions

Conclusion

Novelty and advantages of McRFIS-MN at a glance

- The use of **Memory Neurons** throughout the network at a cellular level helps in capturing the **input-output dynamical relationship** closely
- The **projection-based one-shot learning** is **fast and accurate**
- **Meta-cognitive** structure learning is **highly effective**

Future directions of Research

- Incorporation of **type 2 fuzzy inference** into McRFIS-MN for **better handling of uncertainty** in real world time series problems.
- Application of McFIS-MN in **classification problems**

References

[1] T Takagi, M Sugeno, “Fuzzy identification of systems and its applications to modeling and control” ,IEEE Cybernatics

[2] H.-J. Rong et al., “Sequential adaptive fuzzy inference system (safis) for nonlinear system identification and prediction” ,Fuzzy Sets and Systems-Elsevier, 2006

[3] K Subramanian et al., “A metacognitive neuro-fuzzy inference system (McFIS) for sequential classification problems”,IEEE TFS,2013

[4] M Pratama, SG Anavatti, PP Angelov, “PANFIS: A novel incremental learning machine” , IEEE TNN,2014

[5] P. J. Werbos, “Backpropagation through time: what it does and how to do it”, Proceedings of the IEEE, 1990

[6] GS Babu, S Suresh ,“Meta-cognitive RBF network and its projection based learning algorithm for classification problems”, Applied Soft Computing, 2013 – Elsevier

[7] A. U. Levin et al. , “Control of nonlinear dynamical systems using neural networks. ii. observability, identification, and control" , IEEE TNN,1996

[8] P. S. Sastry, G. Santharam et al., “Memory neuron networks for identification and control of dynamical systems”, IEEE Trans. Neural Networks ,1994

[9] K. S. Narendra and K. Parthasarathy, “Identification and control of dynamical systems using neural networks”,IEEE TNN,1990

[10] R. S. Crowder, “Predicting the Mackey-Glass time series with cascade-correlation learning”,Elsevier,1991

[11] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, “Time series analysis: forecasting and control” John Wiley & Sons, 2015


THANK
YOU