

TECHNICAL CONTRIBUTIONS [DRAFT]

**TIME SERIES FORECASTING OF MONTHLY ACTIVE
USERS WITH LONG SHORT-TERM MEMORY NETWORKS**

Matthew Louis Rosendin
University of California, Berkeley
Department of Industrial Engineering and Operations Research

April 13, 2018

Contents

1	Introduction	3
1.1	Motivation	3
1.2	Goal	3
2	Object-Oriented Abstraction	3
2.1	Decomposition	3
2.2	Encapsulation	3
3	Hyperparameter Optimization	3
3.1	Cross-validation	3
3.2	Grid Search	3
3.3	Results	3
4	Systems Engineering	3
4.1	Technology Stack	3
4.2	Visualizations	3
4.3	Job Scheduling	3
4.4	Documentation	3
5	Next Steps	3
5.1	Data Pipeline Automation (Apache Hadoop)	3
5.2	Dimensionality Reduction by Feature Selection	3
5.3	Improving Hyperparameter Optimization Efficiency	3
5.4	Distributed Machine Learning Clusters	3
6	Conclusion	3
7	Reference	3

1 Introduction

1.1 Motivation

1.2 Goal

2 Object-Oriented Abstraction

2.1 Decomposition

2.2 Encapsulation

3 Hyperparameter Optimization

3.1 Cross-validation

3.2 Grid Search

3.3 Results

4 Systems Engineering

4.1 Technology Stack

4.2 Visualizations

4.3 Job Scheduling

4.4 Documentation

5 Next Steps

5.1 Data Pipeline Automation (Apache Hadoop)

5.2 Dimensionality Reduction by Feature Selection

5.3 Improving Hyperparameter Optimization Efficiency

5.4 Distributed Machine Learning Clusters

6 Conclusion

7 Reference