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

1.1 Motivation

1.2 Goal

Some goals: - Automate model - Read products from dataset automitically -

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   Server Infrastructure

4.2   Distributed Task Queue

4.3   Forecast Visualization

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   Statistical Significance Boundary Visualization

5.5   Distributed Machine Learning Clusters

6   Conclusion

7   Reference