**Казеев Никита Александрович Применение методов машинного обучения к идентификации частиц в детекторе LHCb**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Казеев Никита Александрович

Contents

Chapter 1 Introduction

1.1 New Physics and the LHCb Experiment in Search of It

1.2 Machine Learning

1.3 My Contribution

Chapter 2 Machine Learning

2.1 A Very Brief History of Artificial Intelligence

2.2 Machine Learning Formalism

2.2.1 Model and Training

2.2.2 Hyperparameters

2.3 Measuring Model Quality

2.3.1 Accuracy

2.3.2 Mean Squared Error (MSE)

2.3.3 LogLoss

2.3.4 Area Under the Receiver Operating Characteristic (ROC AUC)

2.4 No Free Lunch Theorem

2.4.1 Formalism

2.4.2 Example

2.4.3 Implications

2.5 Deep Learning

2.5.1 Logistic Regression

2.5.2 Deep Neural Networks

2.5.3 Optimisation

2.5.4 Training Deep Neural Networks

2.5.5 Designing Neural Networks

2.5.6 Implementing Neural Networks

2.5.7 Conclusion

2.6 Gradient Boosting Decision Tree (GBDT)

2.6.1 Decision Tree

2.6.2 Boosting

2.6.3 Implementing GBDT

2.6.4 Conclusion

2.7 Generative models

2.7.1 Generative Adversarial Network (GAN)

2.7.2 Wasserstein GAN

2.7.3 Cramer (Energy) GAN

2.8 Conclusion

Chapter 3 Machine Learning in High-Energy Physics

3.1 Training and Validation

3.2 HEP-specific Machine Learning

3.2.1 Learning to Pivot with Adversarial Networks

3.2.2 Boosting to Uniformity

3.3 Primary Applications

3.3.1 Event Selection: Separating Signal and Background

3.3.2 Event Reconstruction

3.3.3 Monitoring and Data Quality

3.4 Conclusion and Outlook

Chapter 4 The LHCb experiment

4.1 The Large Hadron Collider (LHC)

4.1.1 The LHC Accelerator System

4.1.2 The Large Experiments at the LHC

4.2 The LHCb Detector

4.2.1 Tracking

4.2.2 Particle Identification

4.3 LHCb Data Processing

4.3.1 Hardware Trigger (L0)

4.3.2 Software Trigger (HLT)

4.3.3 Offline Data Processing

4.3.4 Historical Perspective: Run

4.3.5 Upgrade Towards Run

4.3.6 HLT1 on GPU (Allen)

4.3.7 Calibration Samples

4.3.8 Machine Learning at LHCb

Chapter 5 Muon Identification

5.1 Muon Detector

5.2 muDLL

5.3 Correlated x2

5.4 Machine learning for Run II

5.5 Machine Learning Towards Run III

5.6 Algorithms Evaluation

5.7 Data Analysis Olympiad (IDAO)

5.7.1 Introduction

5.7.2 Muon ID Competition

5.8 Conclusion

Chapter 6 Machine Learning on Data With sPlot Background

Subtraction

6.1 sPlot

6.2 The Problem of Negative Weights

6.3 Related Work

6.4 Proposed Approaches

6.4.1 sWeights Averaging (Constrained MSE)

6.4.2 Exact Maximum Likelihood

6.4.3 Classes with Separate Background

6.5 Experimental Evaluation

6.5.1 UCI Higgs

6.5.2 LHCb Muon Identification

6.6 Conclusion

Chapter 7 Global Charged Particle Identification

7.1 Objective and Formalisation of the Global PID

7.2 Adding Likelihoods

7.3 Combining Information with Machine Learning

7.4 State-of-the-art Machine Learning

7.5 Performance

7.5.1 Simulation

7.5.2 Real Data: Calibration Samples

7.6 Conclusion

Chapter 8 Fast Simulation of the Cherenkov Detector

8.1 The Role of Simulated Data in High-Energy Physics Experiments

8.1.1 Detector Design

8.1.2 Data Analysis

8.2 Simulation in LHCb

8.2.1 Technical Improvements to Full Simulation

8.3 Fast Simulation

8.3.1 ReDecay

8.3.2 Parametrisation and Simplification

8.3.3 CaloGAN

8.4 Pilot study: BaBar DIRC

8.4.1 DIRC detector

8.4.2 Our model

8.4.3 Evaluation Results

8.5 Fast Parametric Simulation at LHCb (Lamarr)

8.5.1 RICH Fast Simulation

8.5.2 Preliminary Evaluations

8.5.3 Future outlook

8.6 Conclusion and outlook

Chapter 9 Conclusion

Appendix A No Free Lunch Theorem Proof

Appendix B Global PID input variables

B.1 Used in ProbNN and our models

B.2 Additional engineered features

Bibliography