Modern Methods in Data Analysis

Multivariate Data Analysis: Higgs Challenge

Lukas Fritz	1686473
Fabian Leven	1638446
Ali Deniz Özdemir	1724032
Lena Salfenmoser	1723697
Johannes Heizmann	1725035

4. Februar 2015

Inhaltsverzeichnis

1 Introduction **3**

2 Auswahl einer geeigneten Untermenge an Parametern **3**

3 Getting familiar with the project **3**
3.1 Correlation of Variables 3
3.2 Choosing a Classifier 5

4 improvement approach/Methodik **5**
4.1 Improving the Classifier 5
4.2 Choosing the right cut 5

5 Conclusion **5**

1 Introduction

[?]

2 Auswahl einer geeigneten Untermenge an Parametern

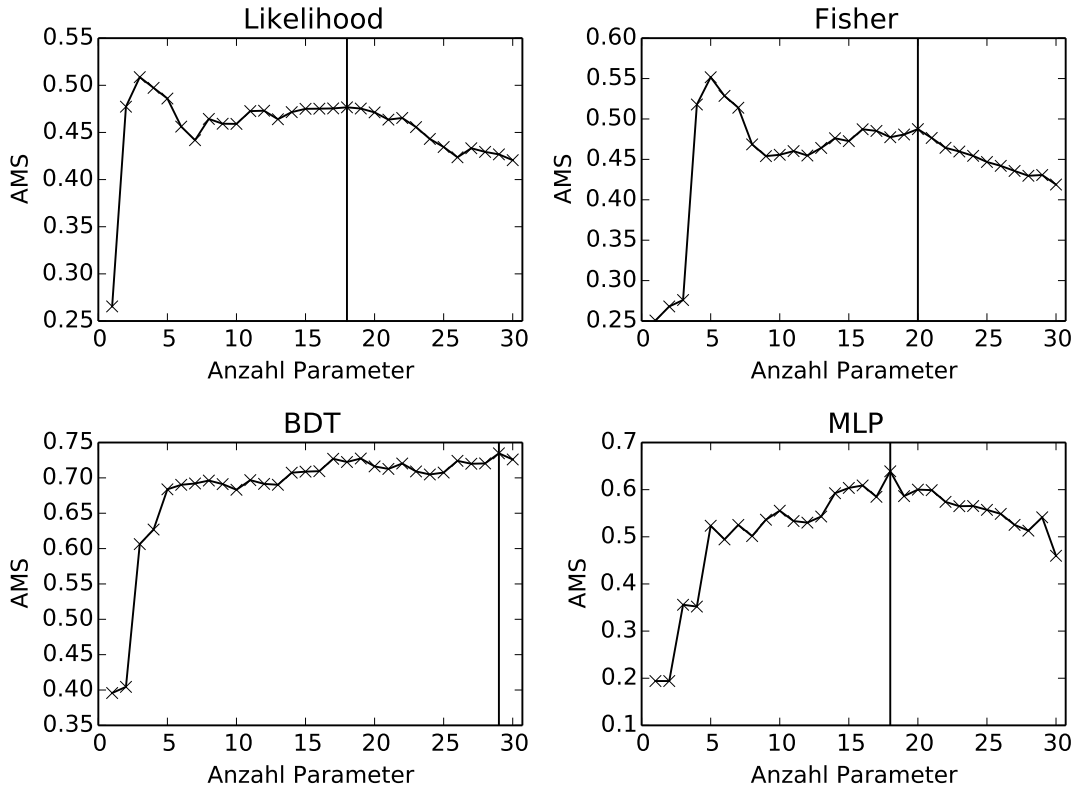

Abbildung 1: AMS über der Anzahl der Parameter.

3 Getting familiar with the project

3.1 Correlation of Variables

- determination of differences in correlations of signal and background
- removal of non relevant variables

Tabelle 1: Bewertung der Parameter. Für die einzelnen Methoden nimmt die Gewichtung der Parameter von oben nach unten ab. Für jede Methode ist gekennzeichnet, welche Parameter nicht mehr verwendet werden.

Likelihood	Fisher	BDT	MLP
d_deltar_tau_lep	d_mass_vis	d_mass_transverse_met_lep	p_jet_num
d_mass_MMC	d_deltar_tau_lep	d_deltar_tau_lep	d_deltaeta_jet_jet
d_pt_ratio_lep_tau	d_mass_transverse_met_lep	d_mass_MMC	d_mass_MMC
d_mass_transverse_met_lep	d_pt_ratio_lep_tau	d_mass_vis	p_jet_leading_pt
d_mass_vis	p_met_phi	d_pt_ratio_lep_tau	d_mass_transverse_met_lep
d_mass_jet_jet	p_jet_num	p_lep_pt	p_jet_leading_eta
d_prodeta_jet_jet	d_mass_jet_jet	d_pt_h	d_mass_jet_jet
d_deltaeta_jet_jet	d_mass_MMC	p_jet_subleading_eta	p_jet_leading_phi
d_lep_eta_centrality	p_jet_subleading_pt	p_lep_phi	d_pt_h
p_lep_eta	d_deltaeta_jet_jet	d_mass_jet_jet	d_mass_vis
d_met_phi_centrality	d_sum_pt	d_lep_eta_centrality	p_jet_subleading_phi
p_jet_subleading_eta	p_tau_pt	p_lep_eta	d_prodeta_jet_jet
p_tau_eta	p_lep_pt	p_jet_subleading_pt	p_jet_subleading_pt
d_pt_h	p_jet_all_pt	d_sum_pt	p_tau_pt
p_tau_phi	p_lep_phi	p_jet_subleading_phi	p_met_phi
p_tau_pt	d_pt_h	p_jet_leading_phi	p_jet_all_pt
p_jet_subleading_pt	p_lep_eta	p_jet_leading_eta	p_met_sumet
d_sum_pt	p_met	p_met_sumet	p_tau_eta
-			-
p_lep_phi	p_met_sumet	p_jet_leading_pt	p_met
p_lep_pt	p_jet_leading_phi	p_tau_eta	d_met_phi_centrality
	-		
p_met	p_jet_leading_eta	d_prodeta_jet_jet	d_pt_ratio_lep_tau
p_jet_leading_pt	d_lep_eta_centrality	p_jet_all_pt	p_tau_phi
p_met_phi	p_jet_subleading_phi	p_jet_num	p_jet_subleading_eta
p_met_sumet	p_tau_eta	p_tau_pt	p_lep_phi
p_jet_subleading_phi	p_jet_subleading_eta	p_met_phi	d_pt_tot
p_jet_num	p_jet_leading_pt	d_deltaeta_jet_jet	p_lep_eta
d_pt_tot	d_pt_tot	p_met	d_sum_pt
p_jet_all_pt	d_prodeta_jet_jet	p_tau_phi	d_lep_eta_centrality
p_jet_leading_phi	p_tau_phi	d_met_phi_centrality	p_lep_pt
		-	
p_jet_leading_eta	d_met_phi_centrality	d_pt_tot	d_deltar_tau_lep

3.2 Choosing a Classifier

4 improvement approach/Methodik

4.1 Improving the Classifier

4.2 Choosing the right cut

5 Conclusion