CDF/PHYS/EXOTIC/CDFR/XXXX
April 26, 2007

Comparison of the NN and ME H$\rightarrow W^+W^-$ Analysis[1]

D. Benjamin[1], P. Bussey[2], D. Hidas[1], S. Hsu[3],M. Kruse[1]
E. Lipeless[3],M. Neubauer[3], A. Robson[2],
R. St. Denis[2], A.S. Thompson[2], F. Wuerthwein[3],W. Yao[5]
[1]Duke University
[2]Glasgow University
[3]University of California- San Diego
[4]University of Illinois- Urbana-Champaign
[5]Lawrence Berekely Laboratories
April 26, 2007
V1.0

Abstract

Comparison of the analyses for H$\rightarrow W^+W^-$based on 1 fb^{-1} of integrated luminosity is made. Approximately 70-80% of the Monte Carlo events containing dibosons selected in the Neural Net analysis are also selected by the Matrix Element analysis, and 63% of the data events are also selected by the Matrix Element analysis. Details on all these overlaps are presented in tables. The is a high degree of correlation between the Matrix Element Likelihood Ratios and the Neural Net scores. However for WW events, although there is a large contribution for very low scores, when the score is above 0.1, there is almost no correlation between the NN and ME scores.

1 Introduction

This note describes a comparison of the analyses of 1fb^{-1} of data in looking for a contribution of H$\rightarrow W^+W^-$ to high p_t dileptons. The Matrix Element

[1]Hww/doc/CompareBlessed1fb/CompareBlessed.tex

analysis uses likelihood ratios as described in [1] and the neural net analysis described in [2] trains artificial neural nets to distinguish the Higgs contribution. The two scores are compared to one another for the common sample of Data and Monte Carlo events described in the corresponding references.

2 Analysis

The analysis was conducted on files that were produced from the Matrix Element (ME) and the Neural Net analyses (NN). Only the LR and NN discriminants for a Higgs mass, $M_h = 160$ GeV/c^2 were considered. All inputs event to these analysis came from the same Stntuples. The lepton identification and event selections differed. Of particular note is the fact that the ME analysis did not use events with jets. The NN analysis used a Drell-Yan neural net to reject Drell-Yan events and the ME analysis used a missing energy variable, "MetSigSpecial". In the NN analysis scores were computed for all events; however, due to the time that it takes to compute the ME for the Likelihood Ratio (LR), only a subsample of events in Monte Carlo have LR computations. The lists of events for the ME analysis consists then of two categories. First, there are events that are selected for the final plot of the ME values and have LR computations available. This is called the ME-LR sample. Next there are events that would have been selected but for which no LR computations have been done. This is called the ME-NoLR sample.

In comparing the analysis, the events in both the ME-LR and ME-NoLR samples are compared to those in the NN sample. This gives the overlap of events since the inputs are the same. Then events in the ME-LR sample that overlap with those in the NN sample are studied and six plots are presented:

- ME-LR vs NN Score (box style)
- ME-LR vs NN Score (lego style)
- ME-LR vs NN Score (box style, high scores)
- ME-LR vs NN Score (box style, low scores)
- NN Score - ME-LR vs NNScore
- NN Score - ME-LR vs NNScore + ME-LR

3 results

Table 1 shows the overlap of events and the number compared. The WW sample was only done for Pythia as the NN analysis has not been done for MC@NLO. For the ME analysis, the $W\gamma$ sample is treated separately for W decays to e, μ and τ, whereas for the NN analysis only the electron and muon decays were analyzed and were not separated.

For Higgs Monte Carlo events selected by the Matrix Element analysis, 46% overlap with the Neural Net analysis. The WW, WZ and ZZ events have about 30% of the Matrix Element analysis events overlap with the Neural Net analysis and there is only 5% of the Matrix Element top events overlap with the Neural Net analysis. There is a small overlap in Fakes and the data overlap by about 24%. For Higgs Monte Carlo Events selected by the Neural Net analysis, 72% of the events overlap with the Matrix Element analysis. For WW, ZZ, WZ, and top 83%, 78% , 62% and 42% of the events in the Neural Net analysis overlap with the Matrix Element analysis. For fakes 7% of the events in the Neural Net analysis overlap with the Matrix Element analysis. For the data, 63% of the Neural Net analysis events overlap with the Matrix Element Analysis. There seems to be no overlap in Wγ events.

The plots for each sample of events are shown in Figures 1 to 10. The is a high degree of correlation between the Matrix Element Likelihood Ratios and the Neural Net scores. However for WW events, although there is a large contribution for very low scores, when the score is above 0.1, there is almost no correlation between the NN and ME scores.

There are some mysteries (bugs?):

- Why are there any NoLR Data events?
- Why is there no overlap at all in the Wγ events

There are also some analyses missing from the NN:

- Up and down variations from fakes
- MC@NLO

Events that are in the NN analysis but not in the ME analysis are likely to be due to:

- The lower $M_{\ell\ell}$ cut in the NN analysis

- The looser jet cut

Events that are in the ME analysis but not in the NN analysis are likely to be due to the extra acceptance gained by including more lepton categories.

A number of desirable studies are to be done:

- Study the dilepton categories in the common and different samples.
- Separate the low and high background regions in the ME analysis.
- Determine if the same leptons are actually used in the events that are common.
- Study the NJet dependence of the scores.
- Examine the region outside the peak at low scores in the simluated diboson events.

4 Conclusion

The ME and NN analyses show a large overlap in the scores but there is likeley to be some gain in combining the methods. Understanding why there is any spread in the scores and what the origin is would be quite useful. It is particularly curious that the WW scores that are not very low show very little correlation between the two analyses.

References

[1] CDF 8719

[2] CDF 8700

	HWW	WWP	ZZ	WZ	$t\bar{t}$
Total NN	143133	195811	43727	72331	1215
Total ME	230425	492947	131818	190345	20367
NN not ME	40425	33540	9590	27379	707
ME not NN	127717	330676	97681	145393	19859
Overlap	102708	162271	34137	44952	508
Total ME-LR	8369	9634	10350	7331	9919
ME-LR not NN	4523	6436	7479	5513	9919
Overlap ME-LR	3846	3198	2871	1818	485

	Zee	Zmumu	Ztautau	Weγ	W$\mu\gamma$
Total NN	220	276	478	1977	1977
Total ME	698	1162	454	574	556
NN not ME	187	246	476	1977	1977
ME not NN	665	1132	452	574	556
Overlap	33	30	2	0	0
Total ME-LR	349	580	227	287	278
ME-LR not NN	316	550	225	287	278
Overlap ME-LR	33	30	2	0	0

	Fakes	Data
Total NN	308	121
Total ME	3694	646
NN not ME	285	45
ME not NN	3671	570
Overlap	23	76
Total ME-LR	1824	323
ME-LR not NN	1847	247
Overlap ME-LR	23	76

Table 1: Comparison of numbers of events in the ME and NN analysis for each process that produces high p_t dileptons.

	HWW	WWP	ZZ	WZ	$t\bar{t}$
NN not ME	28	17	22	38	58
ME not NN	55	67	74	76	98
Overlap/NN	72	83	78	62	42
Overlap/ME	45	33	26	24	2
ME-LR not NN	54	67	72	75	100
Overlap ME-LR	46	33	28	25	05

	Zee	Zmumu	Ztautau	Weγ	W$\mu\gamma$
NN not ME	85	89	100	100	100
ME not NN	95	97	100	100	100
Overlap/NN	15	11	0	0	0
Overlap/ME	5	3	0	0	0
ME-LR not NN	91	95	99	100	100
Overlap ME-LR	9	5	1	0	0

	Fakes	Data
NN not ME	93	37
ME not NN	99	88
Overlap/NN	7	63
Overlap/ME	1	12
ME-LR not NN	99	76
Overlap ME-LR	1	24

Table 2: Comparison of overlap fractions (in%) in the ME and NN analysis for each process that produces high p_t dileptons. For NN not ME, the fraction is taken relative to the total in NN. Similarly for ME not NN. For the overlap of the entire sample, the fraction is taken relative to the NN.

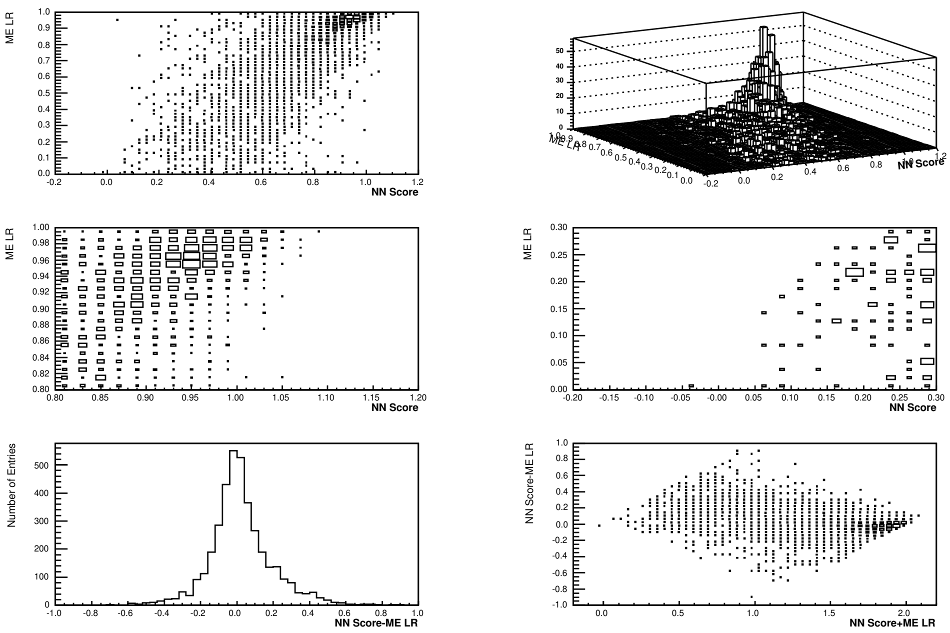

Figure 1: Plots for Higgs at $M_h = 160$ GeV/c^2.

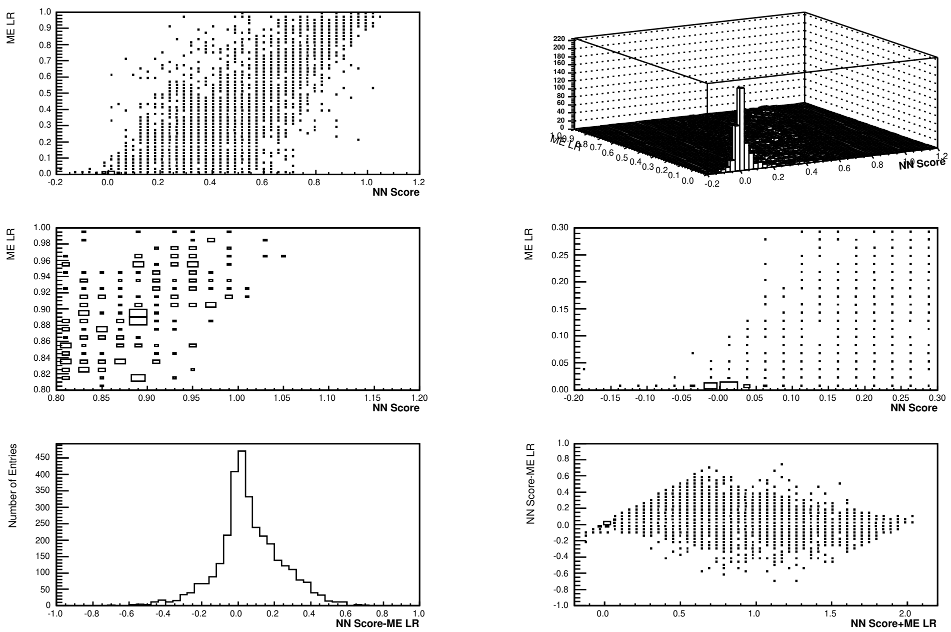

Figure 2: Plots for WW from Pythia based on Higgs at $M_h = 160$ GeV$/c^2$.

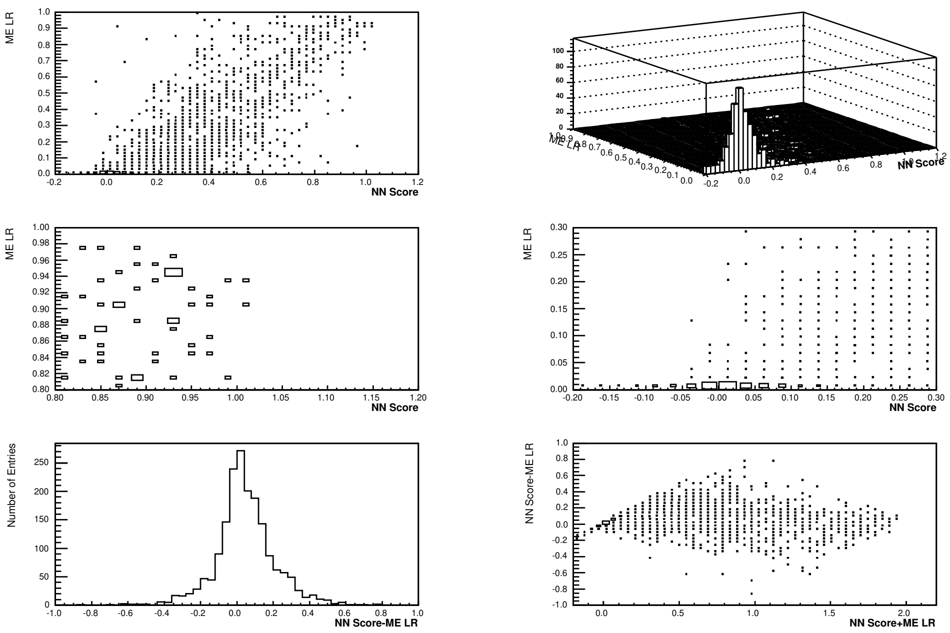

Figure 3: Plots for WZ based on Higgs at $M_h = 160$ GeV$/c^2$.

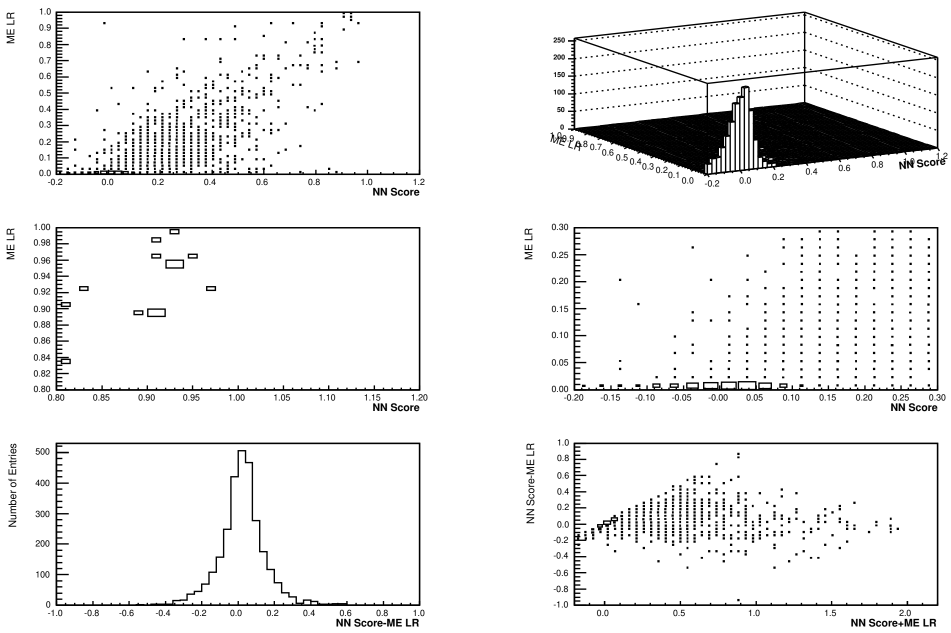

Figure 4: Plots for ZZ based on Higgs at $M_h = 160$ GeV$/c^2$.

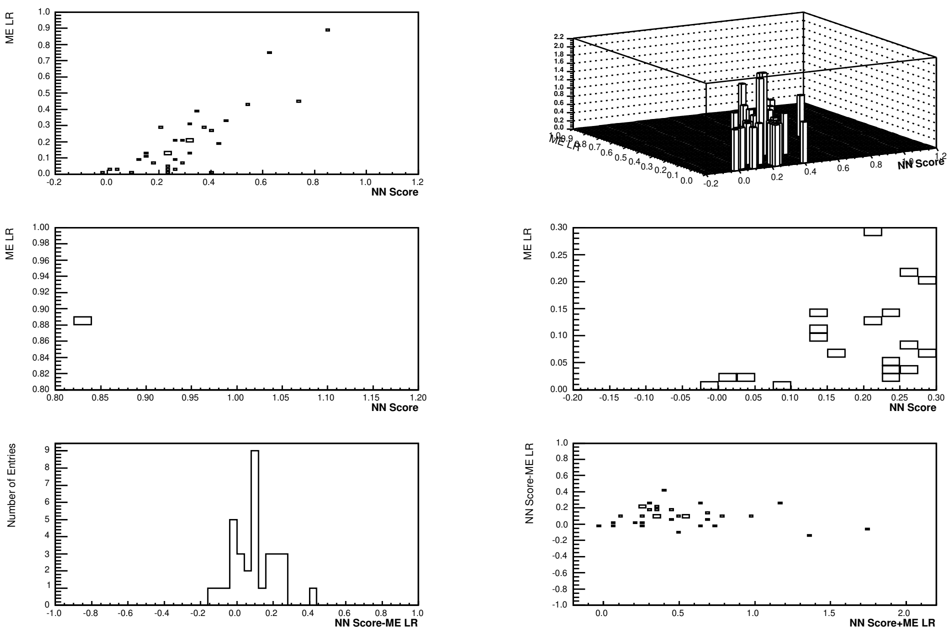

Figure 5: Plots for $Z^0 \to e^+e^-$ based on Higgs at $M_h = 160$ GeV/c^2.

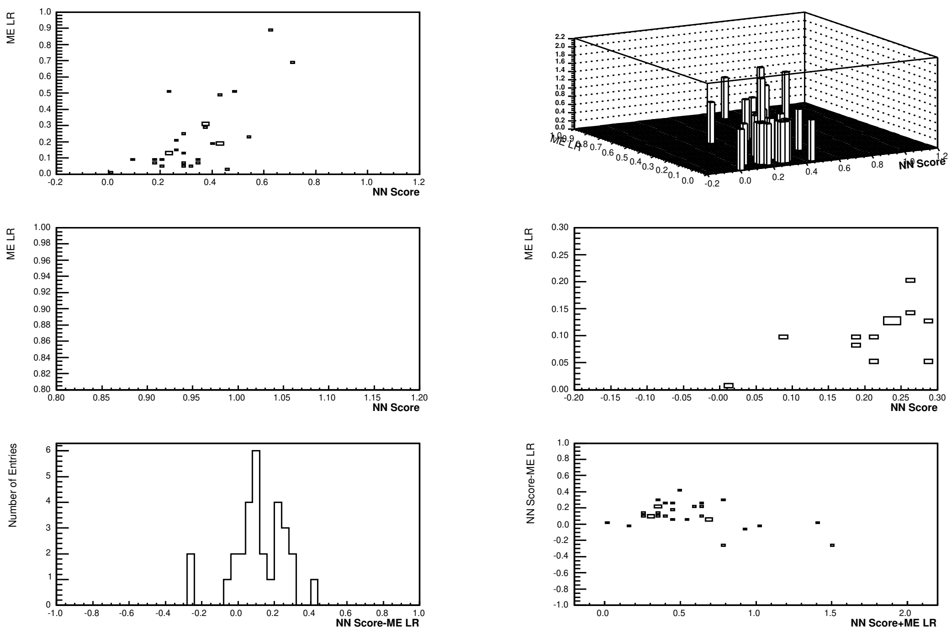

Figure 6: Plots for $Z^0 \rightarrow \mu^+\mu^-$ based on Higgs at $M_h = 160$ GeV$/c^2$.

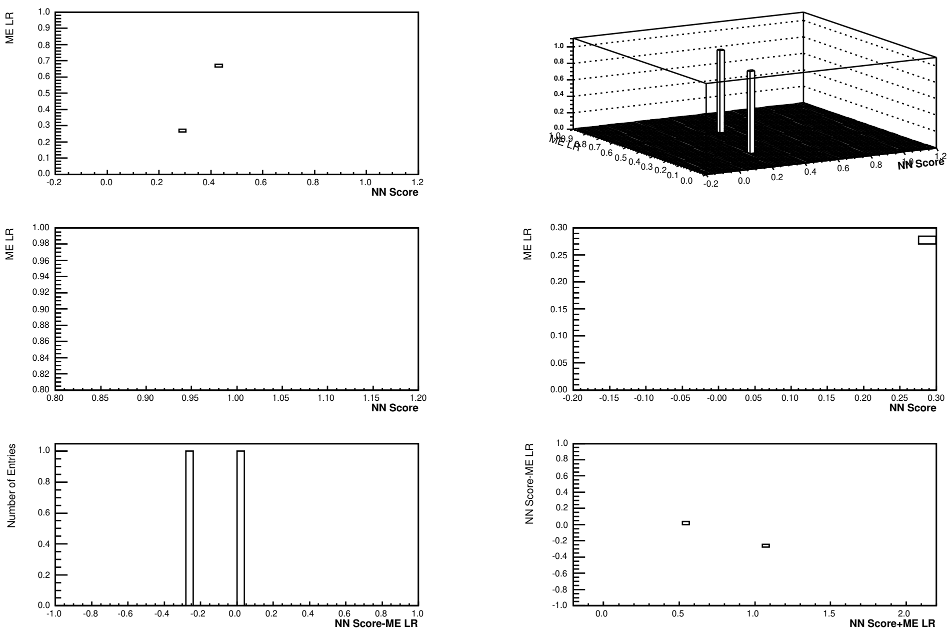

Figure 7: Plots for $Z^0 \to \tau^+\tau^-$ based on Higgs at $M_h = 160$ GeV/c^2.

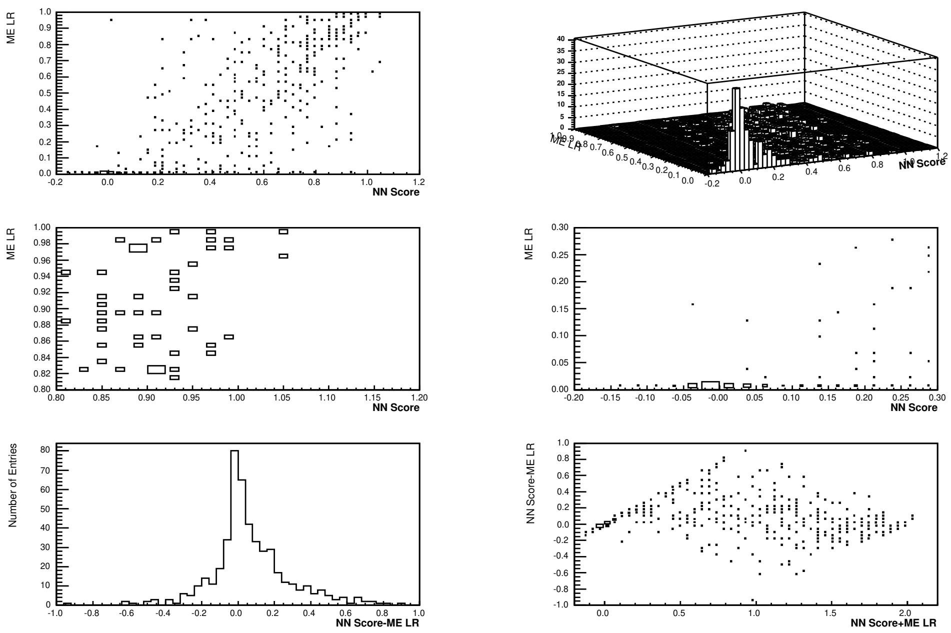

Figure 8: Plots for $t\bar{t}$ based on Higgs at $M_h = 160$ GeV/c^2.

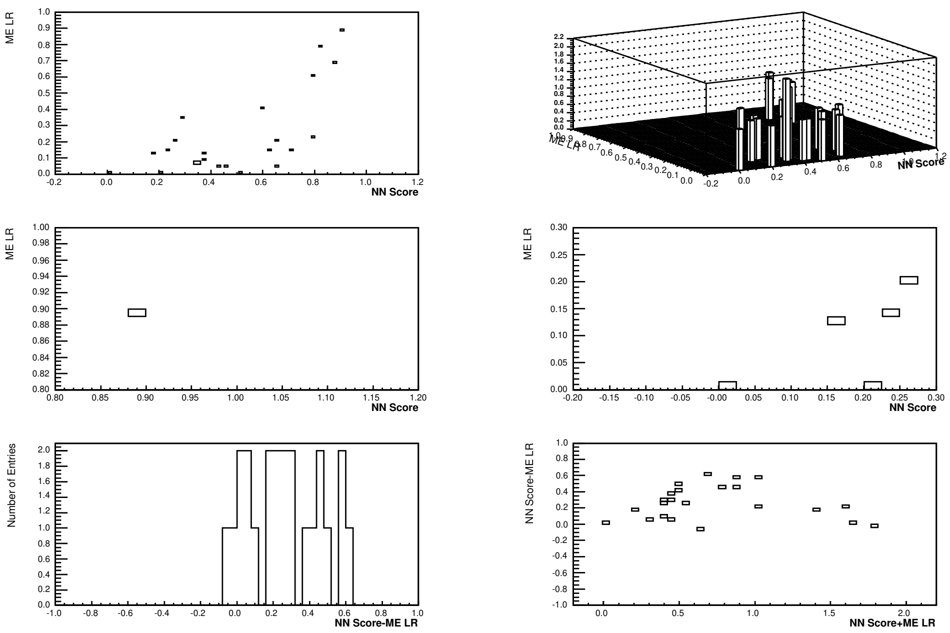

Figure 9: Plots for Fakes based on Higgs at $M_h = 160$ GeV/c^2.

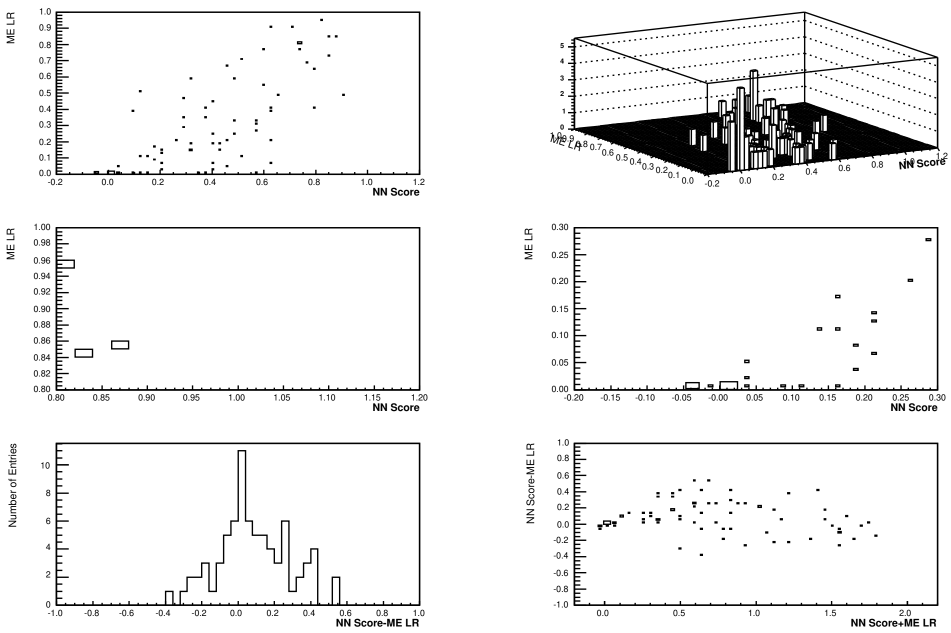

Figure 10: Plots for Data based on Higgs at $M_h = 160$ GeV$/c^2$.