

Создание модели для классификации лептонных распадов W в протон- протонных столкновениях эксперимента ATLAS.

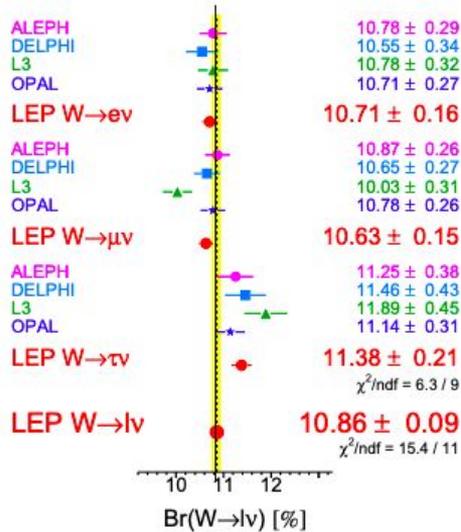
Научный руководитель: Пономаренко Д.Е.

Выполнил: Толкачёв Г.А.

Теоретическое введение

Лептонная универсальность

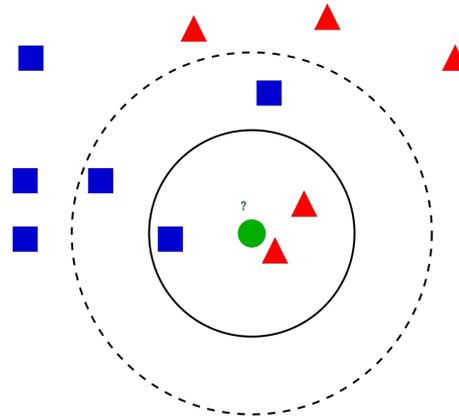
W Leptonic Branching Ratios



В результатах анализа с LEP имеются указания на отклонения от лептонной универсальности в отношениях сечения лептонного распада W в τ к полному сечению лептонного распада W бозона.

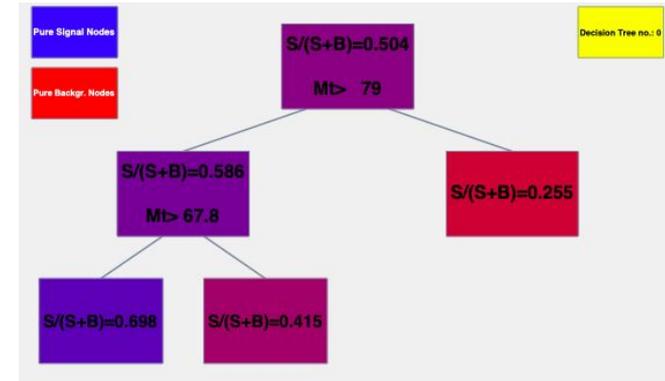
Машинное обучение

Метод k-ближайших соседей (kNN)



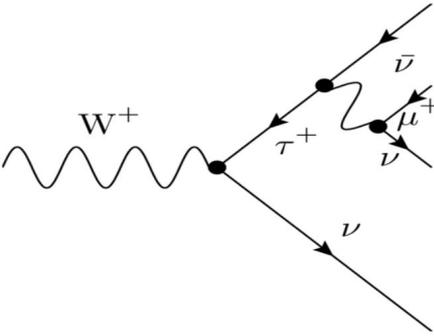
- Вычислить расстояние до каждого из объектов обучающей выборки
- Отобрать k объектов обучающей выборки, расстояние до которых минимально
- Класс классифицируемого объекта это класс, наиболее часто встречающийся среди k ближайших соседей

Метод лес деревьев решений (BDT)



Лес деревьев решений - это множество решающих деревьев. При построении каждого нового дерева учитывает опыт прошлых деревьев таким образом, чтобы оно минимизировало ошибку всех предыдущих деревьев. При завершении обучения мы получаем дерево которое имеет минимальную ошибку при классификации.

Сигнальные и фоновые события



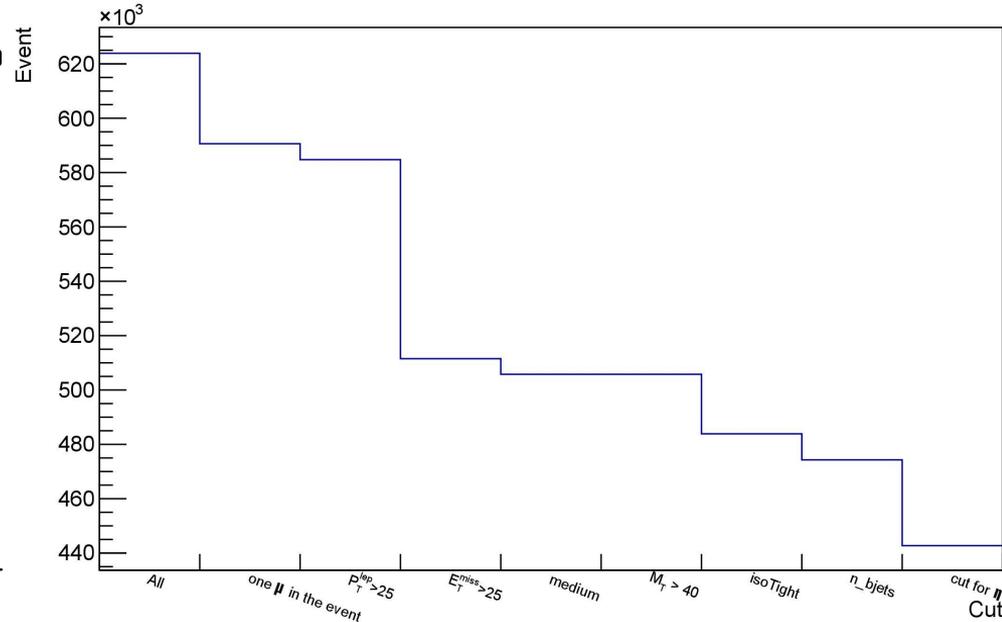
Сигнальные события: Данные используемые в данной работе были получены методом Монте-Карло с помощью генераторов PYTHIA для лептонных распадов W бозона и прошли всю цепочку реконструкций, на условии реальных протонных столкновений эксперимента ATLAS сессия 2 (RUN 2)

$W \rightarrow \tau + \nu \rightarrow \mu + \nu + \nu$
 $W \rightarrow \tau + \nu \rightarrow e + \nu + \nu$

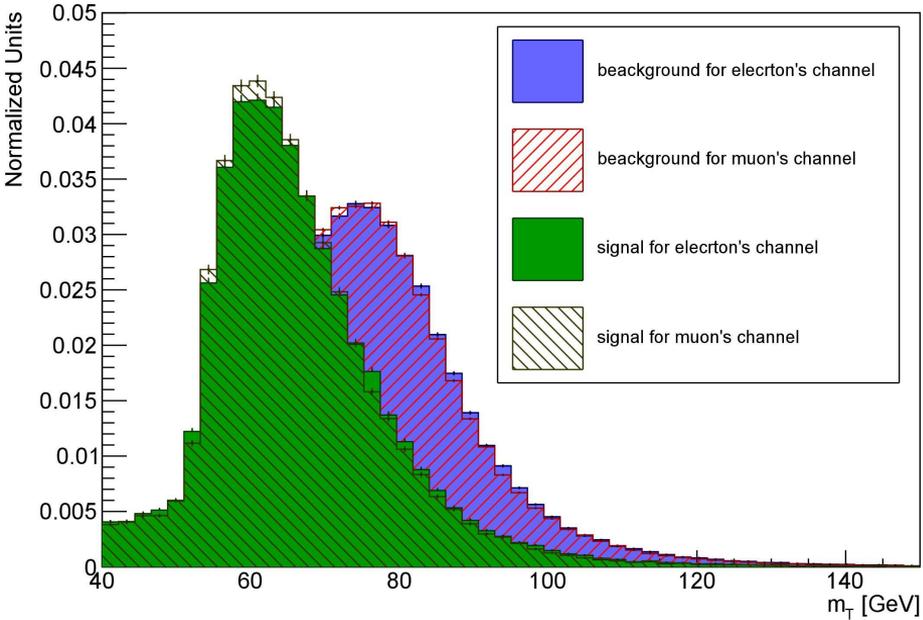
Фоновые события:
 $W \rightarrow e + \nu$
 $W \rightarrow \mu + \nu$

Предварительный отбор
 $E_{\text{miss}} > 25 \text{ ГэВ}$
 $P_t > 25 \text{ ГэВ}$
 $m_T > 40 \text{ ГэВ}$
 $\text{medium} = 1$
 $\text{isoTight} = 1$
 $n_{\text{bjets}} = 0$
Количество лептонов в событии

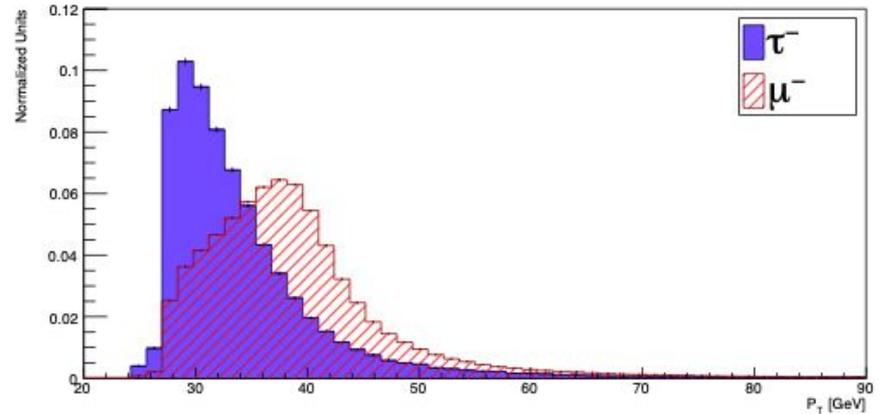
До отбора сигнальные данные содержали 126966 событий а в фоне 623885 .
 После отбора 66779 и 496870.



Сравнение кинематических распределений



Для экономии затрат вычислительных мощностей использован только один канал. Из двух каналов был выбран мюонный, так как он имеет меньше КХД фона чем электронный.



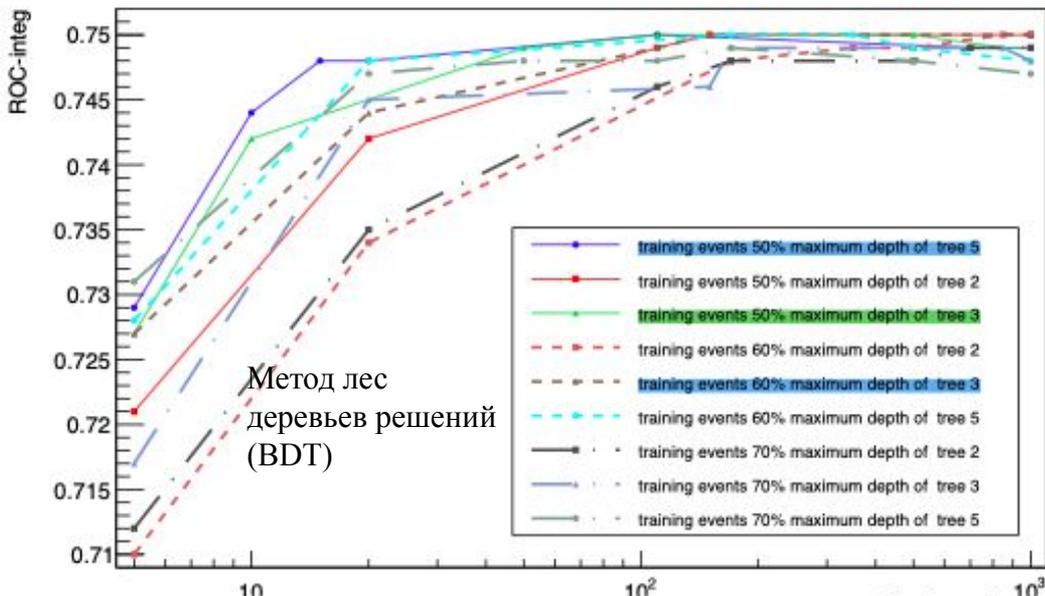
$$m_T = \sqrt{2P_T(l)P_T(\nu)(1 - \cos(\phi(l) - \phi(\nu)))}$$

Исследование методов настроек методов

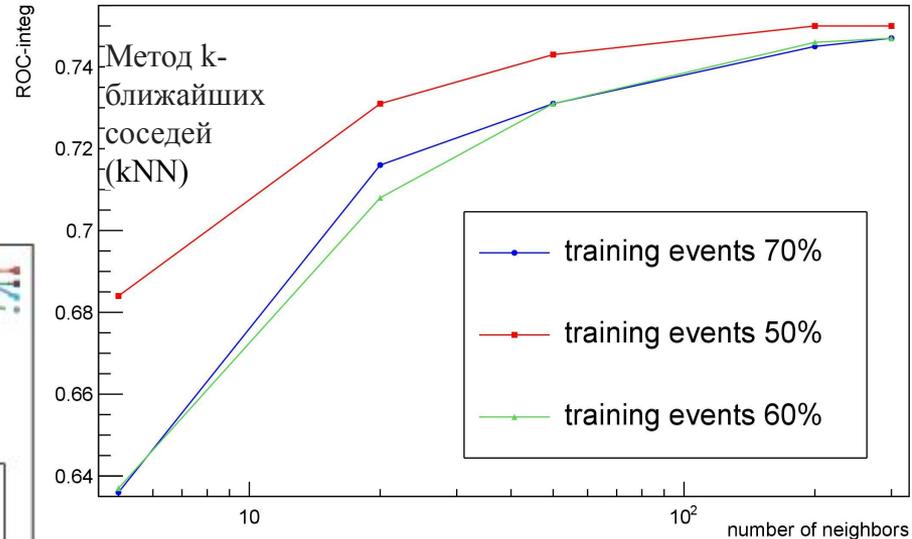
классификации

Кинематические переменные были использованы в качестве входных данных. Наиболее ценный являются:

$$p_T^{lep}, d\Phi(\ell - E_T^{miss}), E_T^{miss}, M_T$$



Сравнение результатов методов



В работе исследовалось влияние конфигурации модели на ее результативность. Были рассмотрены отношение тренировочных и тестовых данных: 50%/50%, 60%/40%, 70%/30%. Для BDT также учитывалось количество деревьев и глубина дерева, а для kNN количество соседей.

Заключение

В процессе выполнения научной работы было выполнено:

- Знакомство с пакетом ROOT , библиотекой машинного обучения TMVA
- Получение начальных сведений о Стандартной модели, эксперименте ATLAS и машинном обучении
- Анализ кинематических распределений для сигнала и фона
- Создание модели для классификации данных
- Исследованы параметры и сделан выбор наилучшей модели для классификации данных методом BDT (50% обучающих данных, максимальная глубина 3, 110 деревьев)

В дальнейшем планируется:

- Использование выбранной модели в анализе лептонной универсальности на эксперименте ATLAS в данных Run 2
- Исследование возможности использования нейронных сетей.
- Знакомство с библиотекой scikit-learn.
- Изучение языка python.

**Спасибо за
внимание!**