

Отчет о педагогической практике

Аспирант группы А19-111

Петухов Александр Максимович

06.06.2023

Введение

В ходе обучения в аспирантуре мной были разработаны прочитаны следующие лекции:

- Метод максимального правдоподобия в физике высоких энергий
- Оценка моделей машинного обучения
- Рекуррентные нейронные сети

Был научным консультантом у двух студентов бакалавриата:

- Константин Савельев
- Алёна Чехонина

Работы студентов

- **Константин Савельев**

Защитил диплом бакалавра по теме «ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ И ФЕНОМЕНОЛОГИЧЕСКИЕ ИЗЫСКАНИЯ ДЛЯ РАЗДЕЛЕНИЯ ЭЛЕКТРОСЛАБОГО И КХД ПРОЦЕССОВ РОЖДЕНИЯ Z-БОЗОНА С ФОТОНОМ»

- **Алёна Чехонина**

Тема работы 5 семестра «ИСПОЛЬЗОВАНИЕ ПЕРЕМЕННЫХ ДОПОЛНИТЕЛЬНЫХ СТРУЙ ДЛЯ ИЗМЕРЕНИЯ ИНТЕГРАЛЬНОГО ПОПЕРЕЧНОГО СЕЧЕНИЯ ПРОЦЕССА АССОЦИИРОВАННОГО РОЖДЕНИЯ ДВУХ Z-БОЗОНОВ В ПРОТОН-ПРОТОННЫХ СТОЛКНОВЕНИЯХ»

Содержание лекций

- **Метод максимального правдоподобия в физике высоких энергий**
 - Оценка числа сигнальных и фоновых событий
 - Моделирование систематических погрешностей
 - Оценка ошибок на параметры
- **Оценка моделей машинного обучения**
 - Критерии оценки
 - Методы разделения данных для оценки моделей
- **Рекуррентные нейронные сети**
 - Работа с последовательностями
 - Рекуррентные нейронные сети (RNN)
 - Сети с длительной краткосрочной памятью (LSTM)

Метод максимального правдоподобия в физике высоких энергий

Цели метода максимального правдоподобия:

1. Оценка теоретических параметров и их ошибок из данных.
2. Проверка гипотез.

Общий подход при использовании статистических моделей:

1. Построить статистическую модель, наиболее полно описывающую физический процесс.
2. Оценить параметры статистической модели, используя статистические методы.
3. Проверить физическую состоятельность полученных оценок.

Метод максимального правдоподобия в физике высоких энергий

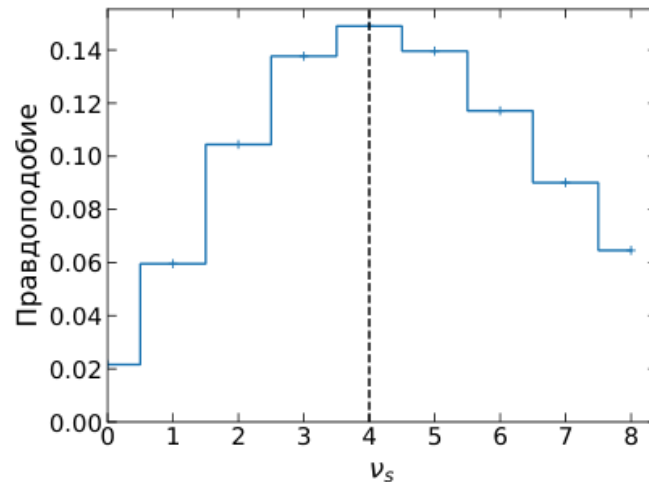
Построение статистических моделей

Счётный эксперимент с известным числом фоновых событий

$$N_{\text{data}} = 7, \tilde{\nu}_b = 3$$

Теперь рассмотрим функцию правдоподобия (фиксированное N_{data} , меняющееся $\nu_s + \tilde{\nu}_b$)

$$L(N_{\text{data}}|\nu_s) = \text{Pois}(N_{\text{data}}|\nu_s + \tilde{\nu}_b) \quad (6)$$



Метод максимального правдоподобия в физике высоких энергий

Нормировочные систематические ошибки

Большинство систематических погрешностей можно моделировать при помощи распределения Гаусса. Данную модель можно рассматривать как калибровочное измерение некоего параметра α' , в результате которого было получено значение a .

$$G(a|\alpha', \sigma_{\alpha'}) = \frac{1}{\sqrt{2\pi\sigma_{\alpha'}^2}} \exp\left[-\frac{(a - \alpha')^2}{2\sigma_{\alpha'}^2}\right] \quad (15)$$

$$a \rightarrow 0$$

$$\alpha' \rightarrow \alpha$$

$$\sigma_{\alpha'} \rightarrow 1$$

$$\eta(\alpha) = \begin{cases} \eta_{+\sigma}, & \alpha = 1, \\ 1, & \alpha = 0, \\ \eta_{-\sigma}, & \alpha = -1, \end{cases} \quad \begin{aligned} L(N_{SR}, 0|\mu, \alpha) &= \text{Pois}(N_{SR}|\mu\tilde{\nu}_s + \eta(\alpha)\tilde{\nu}_b) \cdot G(0|\alpha, 1) \\ &= \text{Pois}(N_{SR}|\mu\tilde{\nu}_s + \eta(\alpha)\tilde{\nu}_b) \cdot \frac{1}{\sqrt{2\pi}} e^{-\alpha^2} \end{aligned}$$

Изменение α на ± 1 — изменение источника систематической погрешности на $\pm\sigma$ и соответствующее изменение числа событий в регионе.

Метод максимального правдоподобия в физике высоких энергий

Погрешности параметров. Отношение правдоподобий

Упрощённый случай: систематики влияют на центральное значение $\hat{\mu}$, но не влияют на её ошибку.

Отношение правдоподобий (likelihood ratio)

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})}; \max[\lambda] = \lambda(\hat{\mu}) = 1.$$

Где $\hat{\mu}$ и $\hat{\theta}$ — оценки параметров, полученные с помощью метода максимального правдоподобия (т.е. при которых $L(\mu, \theta)$ принимает максимальное значение).

Фиксированные значения.

1. Строится статистическая модель и соответствующая ей функция правдоподобия $L(\mu, \theta)$.
2. Ищутся $\hat{\mu}$ и $\hat{\theta}$, при которых $L(\mu, \theta)$ принимает максимальное значение.
3. Функция $\lambda(\mu)$ используется для оценки ошибки $\hat{\mu}$

Оценка моделей машинного обучения

Основная цель машинного обучения — создать модель, дающую точные предсказания на новых данных, т.е. которая умеет **обобщать**.

Также можно встретить термины:

- ▶ **Переобучение** (overtraining/overfitting) — плохая способность модели к обобщению.
- ▶ **Стабильность** — хорошая способность модели к обобщению.

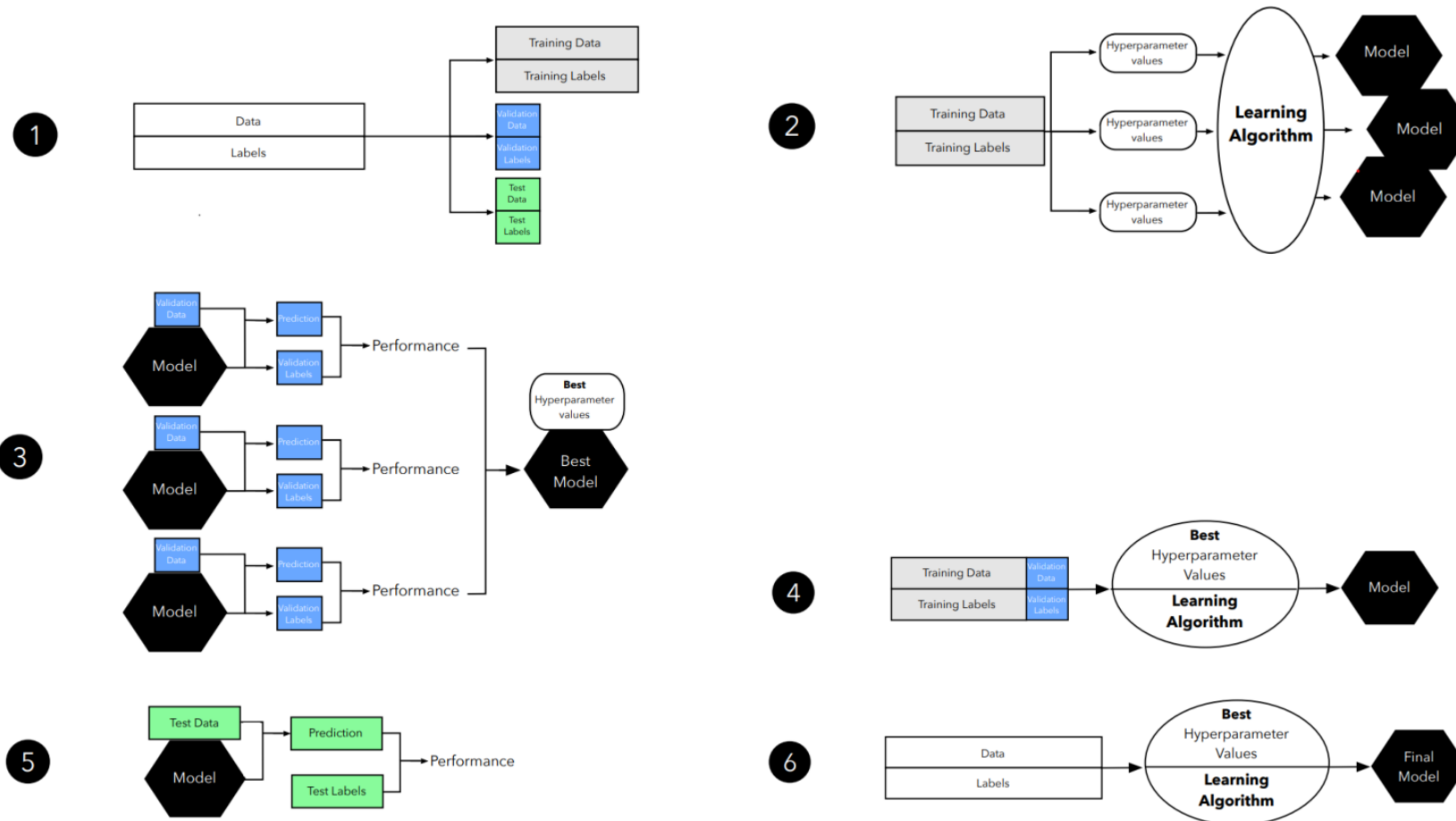
Также известен как компромисс смещения-дисперсии (Bias-variance tradeoff)

$$\text{Ошибка обобщения} = \text{шум} + (\text{смещение})^2 + \text{дисперсия}$$

1. Шум — неустраняемая ошибка, связанная с природой самих данных.
2. Смещение — систематическая ошибка предсказания, общая для всех значений параметров событий.
3. Дисперсия — как сильно изменения параметров событий влияют на отклик модели.

Оценка моделей машинного обучения

Отложенные данные с валидационным набором



Рекуррентные нейронные сети

Проблемы подхода «в лоб»

- Произвольное число элементов – фиксированный размер входного набора
 - Последовательность короче – заполняем остальное заглушками и теряем предсказательную силу
 - Последовательность длиннее – отсекаем полезную информацию
- Каждый элемент может описываться несколькими переменными – входной набор одномерный
 - Обобщать информацию в одно число – теряем предсказательную силу
 - Разворачиваем в комбинацию вектором – значительный рост числа входных переменных
- Информация содержится в относительном расположении элементов друг относительно друга – в простом подходе важна абсолютная позиция переменной
 - Попытка учесть это увеличением числа связей (число скрытых уровней и нейронов) приводит к сложностям оптимизации сетей и склонности к переобучению

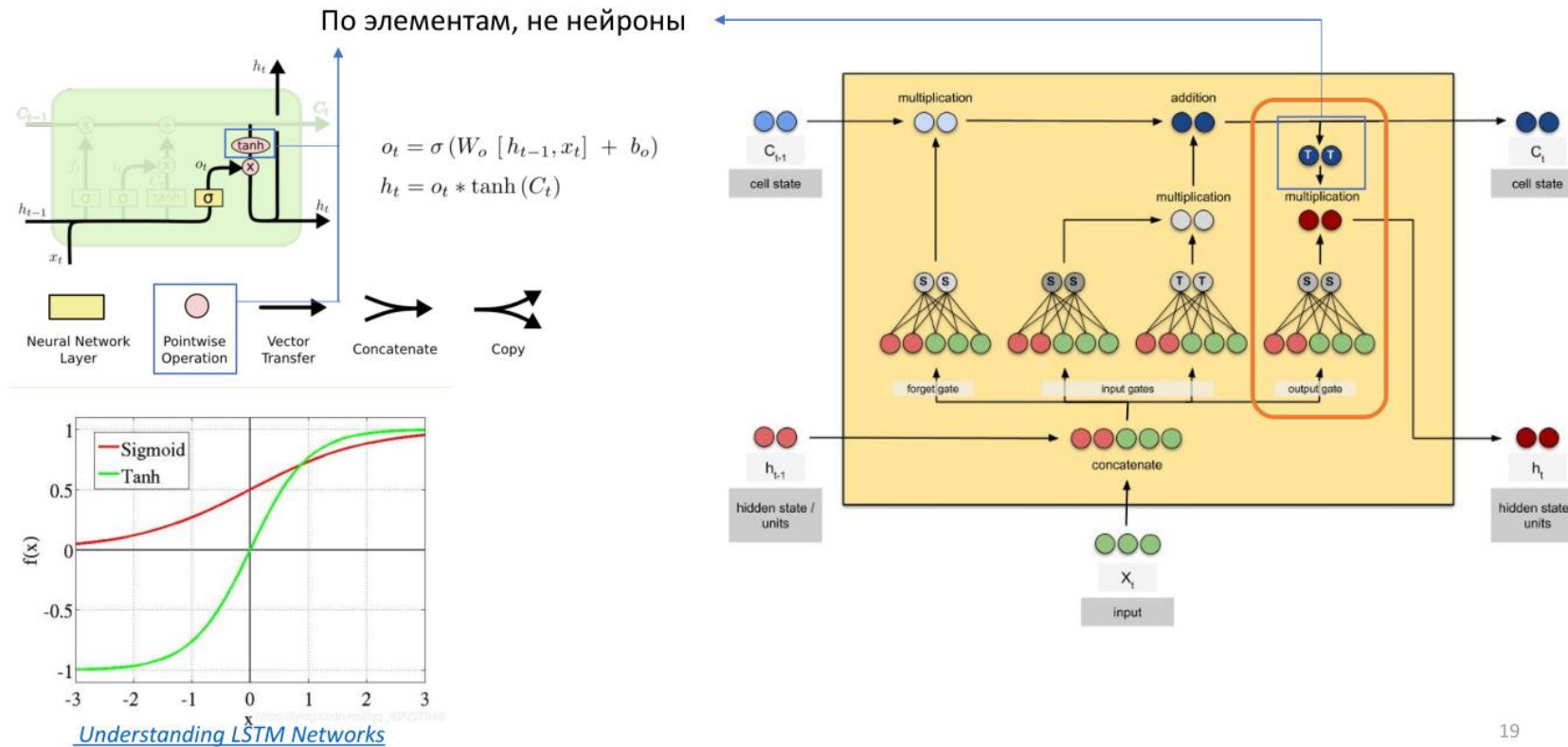
Рекуррентные нейронные сети

Недостатки рекуррентных сетей

- Отклик на следующем шаге зависит от откликов на предыдущих
 - + Позволяет описать зависимости между элементами
 - Замедляет прямое и обратное распространение ошибок
- Общие параметры для каждого шага последовательности
 - + Последовательность можно описать меньшим числом параметров
 - К одним и тем же откликам нейронов применяются одни и те же веса, смещения и функции активации – проблема исчезающего или резко растущего градиента

Рекуррентные нейронные сети

LSTM, принцип работы. 4/4



Заключение

В ходе обучения в аспирантуре

- Было разработано 3 новых лекции, охватывающие темы, ранее не затрагиваемые программой кафедры №40.
- На основе разработанных лекций были проведены занятия для студентов бакалавриата и магистратуры.
- Был научным консультантом у двух студентов бакалавриата, один из которых защитил диплом.