

Санкт-Петербургский государственный университет
Факультет прикладной математики – процессов управления

Тюртеева Анна Сергеевна

Выпускная квалификационная работа

Определение погрешностей параметров магнитов на основе машинного обучения для коллайдера NICA

Уровень образования: магистр

Направление 03.04.01 «Прикладные математика и физика»

Основная образовательная программа ВМ.5521.2021 «Математические и информационные технологии»

Научный руководитель:

доктор физико-математических наук, профессор

Овсянников Дмитрий Александрович

Рецензент:

ведущий исследователь НИИЭФА им. Д.В. Ефремова

Казаченко Николай Ильич

Санкт-Петербург

2023

Оглавление

| | |
|--|----|
| Введение..... | 3 |
| Глава 1. Ускорители..... | 5 |
| 1.1. Историческая справка | 5 |
| 1.2. Классификация ускорителей | 9 |
| Глава 2. Проект «Комплекс NICA» | 12 |
| Глава 3. Машинное обучение..... | 21 |
| Глава 4. Генерация данных и обучение моделей..... | 24 |
| 4.1. Генерация данных и формирование датасета..... | 24 |
| 4.2. Построение и обучение моделей..... | 27 |
| 4.2.1. Линейная регрессионная модель | 27 |
| 4.2.2. Полиномиальная регрессионная модель..... | 28 |
| 4.2.3. Многослойный перцептрон..... | 29 |
| Глава 5. Результаты построения моделей..... | 34 |
| 5.1. Вращение вокруг оси движения пучка частиц (DPSI)..... | 34 |
| 5.1.1. Один магнит..... | 35 |
| 5.1.2. Два магнита..... | 36 |
| 5.1.3. Десять магнитов | 37 |
| 5.1.4. Сорок магнитов | 40 |
| 5.2. Смещение магнита в направлении x (DX) | 41 |
| 5.2.1. Два магнита..... | 42 |
| 5.2.2. Десять магнитов | 43 |
| 5.2.3. Сорок магнитов | 48 |
| 5.3. Смещение магнита в направлении y (DY) | 53 |
| 5.3.1. Два магнита..... | 54 |
| 5.3.2. Десять магнитов | 55 |
| 5.3.3. Сорок магнитов | 58 |
| Заключение | 64 |
| Список используемых сокращений..... | 65 |
| Список литературы | 66 |

Введение

Ускорители заряженных частиц, в настоящее время, являются единственным типом технической аппаратуры для экспериментального изучения мельчайших объектов материи, из которой состоит все живое на Земле [1]. Основное развитие ускорительная техника приобрела в последние 50 лет, что делает ее наиболее молодой и прогрессирующей отраслью технических наук.

Одним из видов ускорителей, используемых в настоящее время, является циклический ускоритель. Он используется для получения частиц высоких энергий, ускорение которых происходит по круговой или эллипсоидной траектории. Примером такого ускорителя может служить сверхпроводящий бустерный синхротрон, который выступает в качестве одной из составных частей «Комплекса NICA» [2]. В рамках данного проекта проводятся фундаментальные и прикладные исследования в области ядерной физики.

Как правило, уже на этапе проектирования и строительства ускорителей закладываются определенные рассчитанные заранее параметры, благодаря которым ученые планируют получить необходимый результат при ускорении заряженных частиц. Однако из-за влияния различных факторов, например, внешние условия, некорректная установка элементов в кольцо ускорителя т.д., экспериментальные данные не всегда соответствуют расчетным. В проекте NICA на бустерном синхротроне одним из таких факторов является погрешность магнитных элементов.

Наиболее распространенным способом получения информации на основе анализа уже имеющихся данных является машинное обучение [3]. К такому методу обычно прибегают при наличии задач, решение которых стандартными методами является невозможным или требует слишком много человеческого ресурса без гарантии на успех.

Целью работы является определение погрешностей параметров магнитов на основе машинного обучения для коллайдера NICA.

Для достижения поставленной цели необходимо решить несколько задач:

1. Изучить магнитно-оптическую схему бустера в проекте NICA;
2. Определить основные параметры магнитов и их погрешности, влияющие на траекторию движения заряженных частиц;
3. Рассмотреть различные модели машинного обучения;
4. Выбрать и реализовать алгоритм машинного обучения для определения параметров погрешностей магнитов.

Глава 1. Ускорители

1.1. Историческая справка

Ускоритель заряженных частиц – это устройство, которое предназначено для ускорения пучков заряженных частиц [4]. Энергия сообщается частицам за счет их взаимодействия с электромагнитным полем.

Стоит предположить, что первая необходимость в искусственно ускоренных заряженных частицах появилась в связи с зарождением экспериментальной ядерной физики и опытами Э. Резерфорда по осуществлению искусственного превращения атомного ядра в 1919 году [5].

На тот момент уже существовали высоковольтные рентгеновские трубки, которые вполне можно было бы назвать ускорителями. Однако технических возможностей и энергии α – частиц, естественных радиоактивных изотопов, оказалось мало для бомбардировки более тяжелых ядер. Одной из целей для постройки ускорителей была возможность получения направленного пучка с большой интенсивностью, а также с энергией порядка нескольких мегаэлектронвольт (МэВ).

Первая искусственная ядерная реакция – расщепление ядра лития протонами с энергией 700 кэВ – была совершена сотрудниками Резерфорда в 1931 году. Именно этот год стоит считать началом истории ускорителей. Первые ускорители состояли из генератора высокого напряжения и ускорительной трубки. При различных модификациях данной конструкции удавалось получить энергию 2,5 МэВ, а затем и 3,6 МэВ [5].

Развитие ядерной физики влекло за собой необходимость в энергиях сравнимых с энергией связи нуклона в ядре (в несколько десятков МэВ). Однако метод электростатического ускорения, используемый ранее, оказался непригодным для такого перехода. Поэтому следующим значимым этапом в истории ускорительной техники стоит считать переход на резонансные методы, которые не требуют высоких напряжений [6].

Идея резонансного метода довольно проста. Для получения энергии в 10 МэВ, частицы необходимо ускорять в электрическом поле с разностью потенциалов 10 млн вольт. Такой же энергии можно достигнуть, если 200 раз пройти электрическое поле с разностью потенциалов 50000 вольт. Соответственно необходимо создать несколько ускоряющих промежутков, электрическое поле которых меняется синхронно с движением частицы.

Модели первых резонансных ускорителей были получены в лабораториях Калифорнийского университета в Беркли. Именно там в 1930-1931 годы были созданы циклический ускоритель (циклотрон) и линейный резонансный ускоритель [5]. И если линейные системы не получили должного применения из-за недостаточного развития техники, то циклотрон имеет особое место в развитии ядерной физики. Уже в 1935 году была получена энергия α – частицы 11 МэВ, а в 1938 году при увеличении диаметра циклотрона – 38 МэВ.

В 1940 году в США был запущен циклический индукционный ускоритель, названный бетатрон [7]. Принцип работы основан на бетатронном условии, которое позволяет поддерживать радиус орбиты почти постоянным. Данное свойство являлось важнейшим с практической точки зрения, что позволило ему стать простым источником тормозного излучения. Однако нерезонансный характер ускорения определял максимально достигнутую энергию в 100 МэВ.

К середине 40-х годов ученые пришли к выводу, что энергии частиц, достигаемые на циклотроне ограничены различными релятивистскими эффектами. Ускорители, построенные на нерезонансных методах, ограничены техническими трудностями.

Открытие μ -мезонов и π -мезонов, ответственных за внутриядерные взаимодействия, и становление новой области науки – физики элементарных частиц – вызвало необходимость в энергиях равных нескольким сотням МэВ

(энергия мюона – 106 МэВ, энергия пиона – 140 МэВ). Решение данной проблемы нашел советский физик В.И. Векслер, который сформулировал принцип автофазировки [8], по которому резонансное ускорение может быть продлено до достаточно больших энергий при умеренных требованиях к параметрам ускоряющего поля. Данный принцип был независимо открыт и Э. Мак-Милоном в США. На основе данного принципа были предложены две ускорительных системы: синхротрон – ускоритель с переменным магнитным полем и постоянной частотой ускоряющего поля и фазотрон – прямой потомок циклотрона с постоянным полем и переменной частотой. Выдвинутые идеи достаточно быстро стали претворяться в жизнь и в 1949 году был запущен первый электронный синхротрон, рассчитанный на энергию 280 МэВ, а к 1960 году в мире существовало несколько синхротронов с энергией порядка 1 ГэВ.

В послевоенное время параллельно с циклическими ускорителями развиваются и линейные ускорители. Практически сразу они были разделены на два подкласса – электронные и ионные [9]. В первом наиболее удачно себя продемонстрировал принцип ускорения, основанный на бегущей электромагнитной волне. В СССР крупнейший линейный ускоритель с энергией 1,8 ГэВ находился в Харьковском физико-техническом институте. Следующим шагом в развитии линейных ускорителей протонов является описанная Л. Альварцем система трубок дрейфа. Эта система до сих пор используется для ускорения до средних энергий (менее 200 МэВ), поскольку для больших энергий считается менее эффективной.

Новым этапом развития ускорителей считается появление сильной (жесткой) фокусировки, что позволило уменьшить сечение пучка частиц, а, соответственно, и поперечное сечение магнитов. Это позволило ученым добиться тех же энергий на меньших установках. Экспериментально этот принцип был проверен Р. Вильсоном на электронном синхротроне. К этому поколению относились 3 синхротрона: на 28 ГэВ в ЦЕРНе, на 30 ГэВ в США, на 7 ГэВ в Москве [5]. Принцип сильной фокусировки был распространен и на

другие типы ускорителей, например, циклотроны, что привело к созданию нового класса машин.

В 1953 году состоялась первая международная конференция по устроителям в Европейской организации по ядерным исследованиям (ЦЕРНе), что открыло возможности для обмена информацией между учеными из разных стран, а также обсуждению новые идеи развития. Одной из таких идей является предложение о проведении ядерных реакций на встречных пучках [10]. Изначально речь шла только о столкновении пучков электронов, однако, по мере развития, стало возможно осуществление электронно-позитронных столкновений. Однако в начале развития установок встречались разноречивые мнения о том, будут ли данные типы ускорителей конкурентноспособными по сравнению с ускорителями на неподвижной мишени. Но это не помешало не только продолжить работу над созданием нового типа ускорителей, но и подтолкнуть традиционные ускорители к дальнейшему развитию. Начиная с 1959 года, стали запускаться ускорители на встречных пучках.

Начиная с 60-х годов параллельно развивались многие области ускорительной техники. Происходили изменения в подходе к основным параметрам ускорителей. Возрастало внимание не только к интенсивности пучка, но и к его качеству: угловому разбросу, поляризации и т.д. Приобрёл значение вопрос о возможности одновременной постановки нескольких экспериментов на одной установке. Была успешно решена проблема со сложной магнитной системой изохронного циклотрона для ускорения протонов до релятивистских скоростей. Ускорение тяжелых ионов с достижением энергий порядка нескольких МэВ/нуклон привело и к развитию циклотронов. Возрождались старые идеи, что привело к формированию нового метода ускорения – коллективному, а соответственно, и новому типу ускорителей. В 1967 году в СССР в Серпухове на базе Института физики высоких энергий был запущен протонный синхротрон. В 70-е, благодаря идеям Керста об использовании пучков протонов и Джерарда О'Нила о

возможности применения накопительных колец при создании интенсивных пучков частиц, началась эпоха использования коллайдеров. В 1971 году был построен протонный синхротрон ISR, а в 2012 году на Большом Адронном Коллайдере был открыт Бозон Хиггса [11]. Современные коллайдеры способны разгонять не только элементарные частицы, но и ядра тяжелых металлов.

В основе развития ускорительной техники всегда лежали требования физики высоких энергий и стремление ученых понять из чего же состоит наша Вселенная. Увеличения энергий, скоростей, масштабов ускорительной техники требуют не малые как денежные вложения, так и умственные. Все активнее развивается международное сотрудничество. Крупнейшими организациями, в которых сейчас проводятся подобные исследования – это ЦЕРН – европейская организация по ядерным исследованиям, крупнейшая лаборатория физики высоких энергий и ОИЯИ – Объединенный институт ядерных исследований в Дубне.

1.2. Классификация ускорителей

В современном мире существует большое число типов ускорителей в зависимости от основания классификации (Таблица 1). Ускорители делятся по способу создания ускоряющего электрического поля на:

- ускорители прямого действия. Ускорение частиц происходит в постоянном статическом или квазистатическом электрическом поле.
- индукционные ускорители. Изменяется магнитный поток, вследствие чего возникает вихревое электрическое поле. При этом напряженность электрического поля остается постоянной по направлению.
- высокочастотные (ВЧ) резонансные ускорители. Для ускорения частиц используется электрическая компонента поля электромагнитной волны. Данный метод ускорения

отличается от остальных тем, что постоянно происходит изменение направления ускоряющего поля.

- Ускоритель, основанный на коллективном методе. Для ускорения частиц используется электрическое поле, которое создается другими заряженными частицами.

По форме траектории ускорители можно разделить на 2 типа. Первый – это линейные ускорители, траектория которых представляет собой прямую линию, где пучок частиц однократно проходит ускоряющий участок [12]. Второй – циклический ускоритель. В нем пучок частиц движется по окружности или спирали, несколько раз проходя через одну и ту же ускоряющую конструкцию.

Не стоит забывать, что выбор метода ускорения и конструкции ускорителя зависит от желаемых результатов и типа ускоряемых частиц. Это могут быть электроны, протоны и ионы. Так же могут ускоряться и античастицы (позитроны и антипротоны).

Более того, возможна классификация по характеру ускоряющего поля на резонансный и нерезонансный.

Таблица 1. Классификация ускорителей заряженных частиц

| По типу создания электрического поля | По типу траектории | |
|--|--|---|
| | Линейная | Окружность/спираль |
| <i>Постоянное эл. поле</i> | Ускоритель прямого действия | - |
| <i>Вихревое эл. поле</i> | Линейный индукционный ускоритель (линейный бетатрон) | Бетатрон |
| <i>Электрическая компонента электромагнитной волны</i> | Линейный резонансный ускоритель | Высокочастотные (ВЧ) резонансные ускорители |

| | | | |
|--|--|-------------------------|---|
| | <i>Электрические поля в пучках электронов или плазме</i> | Коллективный ускоритель | - |
|--|--|-------------------------|---|

В таблице 1 все указанные ускорители, кроме линейного резонансного ускорителя и группы ВЧ резонансных ускорителей, являются нерезонансными. В свою очередь циклические высокочастотные резонансные ускорители так же имеют свою классификацию в зависимости от постоянства магнитного поля, изменения радиуса движения пучка частиц, изменения частоты ускоряющего поля и т.д. (Таблица 2) [9,12,13].

Таблица 2. Классификация высокочастотных резонансных ускорителей

| Название ускорителя | Магнитное поле | Постоянство радиуса | Частота ускоряющего поля | Кратность ускорения | Ускоряемые частицы |
|------------------------|----------------|---------------------|--------------------------|---------------------|--------------------|
| Циклотрон | Постоянное | Переменный | Постоянная | Постоянная | Протоны, ионы |
| Микротрон | Постоянное | Переменный | Постоянная | Переменная | Электроны |
| Фазотрон | Постоянное | Переменный | Переменная | Постоянная | Протоны |
| Синхрофазотрон | Переменное | Постоянный | Переменная | Постоянная | Протоны |
| Электронный синхротрон | Переменное | Постоянный | Постоянная | Постоянная | Электроны |
| Ионный синхротрон | Переменное | Постоянный | Постоянная | Постоянная | Ионы |

В приведенных классификациях отсутствуют ускорители на встречных пучках, т.к. пучки от ускорителей, приведенных в таблицах 1 и 2, направляют навстречу друг другу.

В работе изучение погрешности параметров магнитов проводится на основе данных из сверхпроводящего синхротрона.

Глава 2. Проект «Комплекс NICA»

В настоящее время на территории Объединённого института ядерных исследований (ОИЯИ) реализуется проект «Комплекс NICA», нацеленный на экспериментальное исследование ядерной материи, которая возникла во вселенной на первой стадии ее развития [2]. Несмотря на большие достижения в ядерной физике и физике высоких энергий еще многие аспекты окружающего мира, устройства вселенной, процессов, происходящих в микромире, остаются необъясненными.

Воссоздание данных процессов возможны при помощи современных ускорительных комплексов. В мире существует 2 коллайдера, в которых сталкивают пучки частиц с энергиями от сотен до тысяч миллиардов электронвольт. Это способствует изучению процессов зарождения Вселенной и момента возникновения кварков и глюонов. Этими коллайдерами являются Большой адронный коллайдер (LHC) в ЦЕРН и коллайдер Брукхейвенской национальной лаборатории RHIC (США) [14].

В процессе расширения вселенной и ее охлаждения кварк-глюонная материя перешла в более понятную для физиков материю, состоящую из протонов и нейтронов. Эта материя, в свою очередь, может быть воссоздана столкновением тяжелых ионов, но с более низкими энергиями: от нескольких единиц до нескольких десятков миллиардов электронвольт. Для проведения таких исследований и строится ускорительно-экспериментальный комплекс NICA (рис.1) [2,12].

NICA – Nuclotron-based Ion Collider facility- установка на встречных пучках на основе ускорителя нуклотрон. Данный проект реализуется с привлечением специалистов из 18 стран-участниц ОИЯИ. Проект был начат в 2009 году, а в 2016 году был подписан документ между ОИЯИ и Российской Федерацией об участии России в создании данного комплекса. В 2017 году

проект был включен в Европейскую организацию по стратегическому развитию исследовательской инфраструктуры (ESFRI).



Рисунок 1. Комплекс NICA

Основные цели создания комплекса NICA:

1. Фундаментальные исследования в области релятивистской ядерной физики, физики спина частиц, а также в области радиационной биологии.
2. Прикладные исследования. Пучки частиц в большом спектре энергий используются для переработки ядерных отходов АЭС, исследований по радиационной медицине, в материаловедческих исследованиях, а также в исследованиях стойкости микросхем космических аппаратов к воздействию радиации.
3. Создание образовательной программы для проведения исследовательской деятельности и подготовки специалистов различного уровня.

В состав комплекса NICA входят инжекционный комплекс, бустер, нуклотрон, коллайдер, MPD детектор, SPD детектор, установка BM@N [12].

Сверхпроводящий бустерный синхротрон является одной из составных частей комплекса NICA. Во всеобщем понимании, бустер – это промежуточный циклический ускоритель, синхротрон, используемый для

предварительного ускорения и формирования пучка с нужной структурой и параметрами для внедрения (инъекции) в основной ускоритель. В рамках комплекса NICA бустерный синхротрон является инжектором тяжелых ионов в Нуклотрон (основную ускорительную установку комплекса) [14].

В Бустере выполняются такие задачи, как:

- накопление $2 * 10^9$ ионов золота, энергия которых равна энергии, необходимой для инъекции;
- ускорение тяжёлых ионов до энергии 578 МэВ/н;
- быстрый вывод ускоренного пучка для его инъекции в нуклотрон.

Бустер (рис.2) включает в себя следующие системы и подсистемы:

- источники питания;
- высокочастотная ускоряющая система;
- магнитно-оптическую систему;
- система инъекции;
- система вывода пучка.

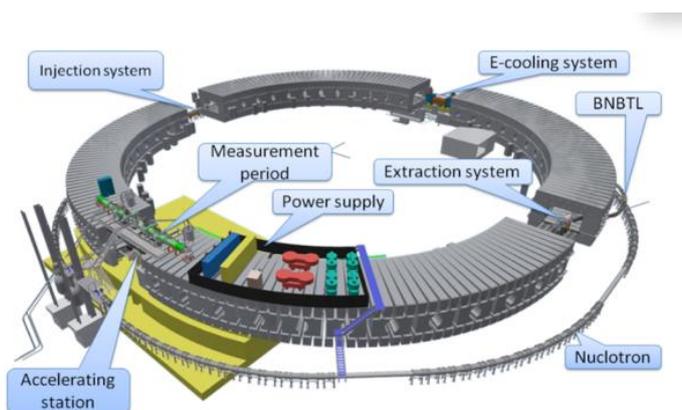


Рисунок 2. Основные структуры Бустера [15]

Подробнее рассмотрим магнитно-оптическую систему, так как содержащиеся в ней элементы задают траекторию движения пучка.

Магнитная структура Бустера состоит из 4-х суперпериодов, в каждый из которых входят 5 регулярных периодов и прямолинейная секция (рис.3).

Суперпериод начинается с прямолинейного участка, после идут регулярные периоды, которые состоят из фокусирующей и дефокусирующей квадрупольных линз, двух дипольных магнитов и четырех малых свободных промежутков, предназначенных для размещения корректоров, коллиматоров и диагностического оборудования (рис.4). Таким образом, каждый суперпериод состоит из десяти дипольных магнитов, шести фокусирующих и шести дефокусирующих квадрупольных магнитов. Прямолинейный участок состоит из тех же элементов, за исключением дипольных магнитов [15].

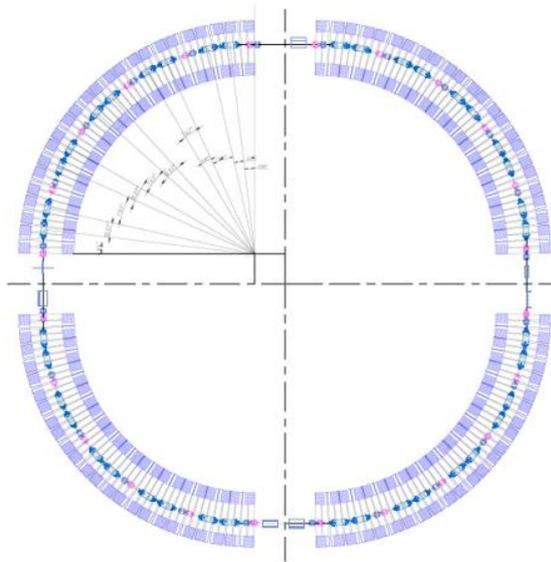


Рисунок 3. Магнитная структура Бустера

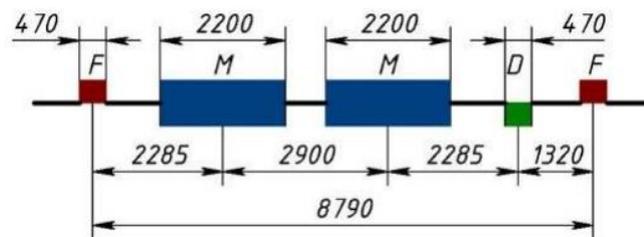


Рисунок 4. Схема регулярного периода. *M* – дипольный магнит, *D* – дефокусирующая квадрупольная линза, *F* – фокусирующая квадрупольная линза

В работе рассматриваются исключительно дипольные магниты. Рассмотренный бустерный ускоритель описывается как последовательность элементов, задающих траекторию пучка, расположенных вдоль опорной орбиты. Опорной орбитой считается путь заряженной частицы, обладающей центральным импульсом, через идеализированные магниты [16]. Поэтому все

перемещения частицы описываются в системе координат, связанной с опорной орбитой.

Данная система координат основана на 3-х нормированных единичных векторах: первый – это вектор касательной к опорной траектории $\vec{t} = \frac{d\vec{r}_0(s)}{ds}$. Здесь мы предполагаем, что $\vec{r}_0(s)$ параметризуется длиной дуги s , что гарантирует, что \vec{t} имеет единичную длину. Вторым единичным вектором \vec{n} называется нормальным вектором и для плоских траекторий без кручения пропорционален скорости ее изменения. $\vec{n} = \frac{-k d\vec{t}}{ds}$, k - кривизна траектории. Знак минус приводит к тому, что \vec{n} указывает в сторону от центра кривизны и гарантирует, что третий единичный вектор \vec{b} , определяемый $\vec{b} = \vec{t} \times \vec{n}$, указывает вверх. Эта конструкция приводит к тому, что три базовых вектора, $\vec{t}, \vec{n}, \vec{b}$ образуют правую систему координат. По общему правилу в литературе эти векторы часто обозначают $\vec{s}, \vec{x}, \vec{y}$. На рис.5 слева показана их ориентация.

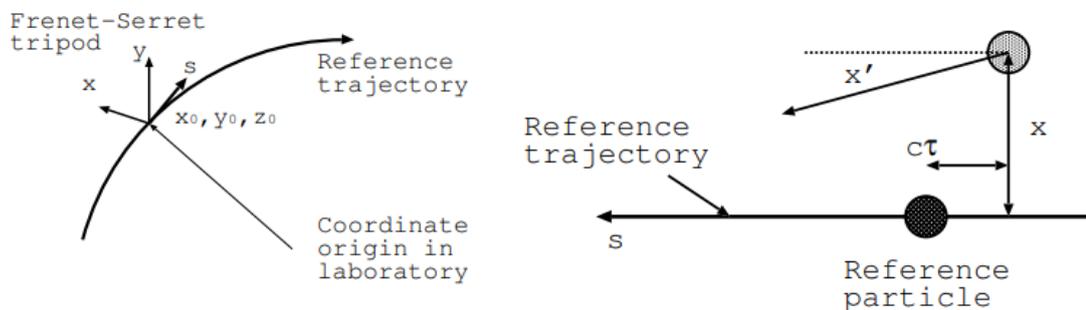


Рисунок 5. Опорная траектория (справа) и совместно перемещающаяся опорная частица, которая используется для определения отклонения (слева) частиц относительно опорной траектории [16]

Преобразуем силы, которые действуют на пучок и определяют уравнение движения к системе отсчёта, связанной с опорной орбитой. Используя сложные математические преобразования, получаем Гамильтониан вида[16]:

$$H = e\Phi + c \sqrt{m^2 c^2 + (p_x - eA_x)^2 + (p_y - eA_y)^2 + \left(\frac{p_s - eA_s}{1 + \frac{x}{\rho}} \right)^2} \quad (1)$$

Здесь Φ и A – потенциалы электрического и магнитного полей, e – заряд, m – масса, p – импульс, ρ – радиус кривизны.

Уравнения Гамильтона в положениях x , y , s и моментах p_x, p_y, p_s дают уравнения движения:

$$\dot{x} = \frac{\partial H}{\partial p_x}, \quad \dot{p}_x = -\frac{\partial H}{\partial x} \quad (2)$$

и аналогично для переменных y и s . Эти уравнения по-прежнему зависят от времени, как от независимой переменной, но использование переменной s в качестве независимой переменной, делает интерпретацию местоположения частиц гораздо более доступной. Мы не хотим знать время, когда частица находится в том или ином положении, а скорее хотим знать их поперечное положение в определенном месте траектории. Тогда Гамильтониан будет иметь вид [16]:

$$G = -\left(1 + \frac{x}{\rho}\right) \sqrt{(1 + \delta)^2 p_0^2 - (p_x - eA_x)^2 - (p_y - eA_y)^2} - eA_s \quad (3)$$

p_0 – импульс эталонной частицы, $\delta = (p - p_0)/p_0$. Кинематическое состояние частицы, для которой требуются три пространственные координаты и три соответствующих импульса или скорости, описывается шестью переменными x, p_x, y, p_y, t и δ .

Основная задача диполей состоит в задании опорной орбиты, но они также оказывают влияние на переменные фазового пространства движения частицы относительно опорной траектории.

Уравнение (3) является основой для последовательного описания уравнений движения, определяющих распространение частиц в ускорителях, и мы можем использовать его независимые переменные для описания

состояния частиц. Поскольку поперечные импульсы p_x и p_y в большинстве ускорителей намного меньше суммарного импульса $p \approx p_0$, мы используем $x' = p_x/p_0$ и $y' = p_y/p_0$. Эти переменные могут быть интерпретированы как углы относительно опорной траектории, что облегчает визуализацию движения частиц, показанного справа на рис.5. Наконец, мы сдвигаем начало временной переменной t , чтобы оно совпадало с началом опорной частицы \hat{t} , $\tau = t - \hat{t}$.

Таким образом, мы приходим к следующим трем позиционным координатам и трем параметрам, связанным с импульсами, которые используются для характеристики частиц в ускорителях.

x - горизонтальное расстояние до опорной орбиты;

x' - горизонтальный угол относительно опорной орбиты;

y - расстояние по вертикали до опорной орбиты;

y' - вертикальный угол относительно опорной орбиты;

τ - время прибытия относительно опорной частицы;

δ - относительная разность импульсов по отношению к опорной частице.

Преимущество этих переменных заключается в том, что они описывают геометрические понятия, такие как расстояния и углы, как показано справа на рис.5. Время прибытия имеет значение, если на частицу воздействуют изменяющиеся во времени электромагнитные поля, например, генерируемые в ускоряющих структурах. Относительную разность импульсов δ удобно использовать, поскольку она описывает относительное отклонение от расчетных углов дипольных магнитов, которое обусловлено отклонением от расчетного импульса частиц. Продольный импульс почти всегда намного больше поперечного, что приводит к тому, что углы x' и y' очень малы.

Все возможные значения кинематических переменных обычно обозначаются как шестимерное фазовое пространство, и определенный набор значений используется для описания состояния частицы. Иногда нет необходимости рассматривать шести фазовое пространство, и достаточно охватить только горизонтальное фазовое пространство, включающее подпространство x и x' , или вертикальное фазовое пространство, включающее y и y' , соответственно. Подпространство, охватываемое τ и δ , обозначается продольным фазовым пространством. Будем ограничиваться этими подпространствами, потому что динамика различных подпространств часто независима.

Параметрами погрешности магнитов являются различные ошибки смещения, интегральные мультипольные коэффициенты и т.д.

В работе рассматриваются ошибки смещения магнитов относительно локальной системы координат. Погрешности определяются как перемещения вдоль осей координат и вращения вокруг этих осей. Влияние смещений рассматриваются в линейном приближении [17].

Ошибки выравнивания определяются параметрами (рис. б):

- DX – смещение в направлении x . По умолчанию смещение 0 м. $DX > 0$ смещает элемент в положительном направлении x .
- DY - смещение в направлении y . $DY > 0$ смещает элемент в положительном направлении y .
- DS - смещение в s -направлении. $DS > 0$ смещает элемент в положительном направлении.
- $DPHI$ - вращение вокруг оси x . По умолчанию: 0 рад.
- $DTHETA$ - вращение вокруг оси y в соответствии с правилом правой руки.
- $DPSI$ - вращение вокруг оси s (оси движения пучка частиц) в соответствии с правилом правой руки.

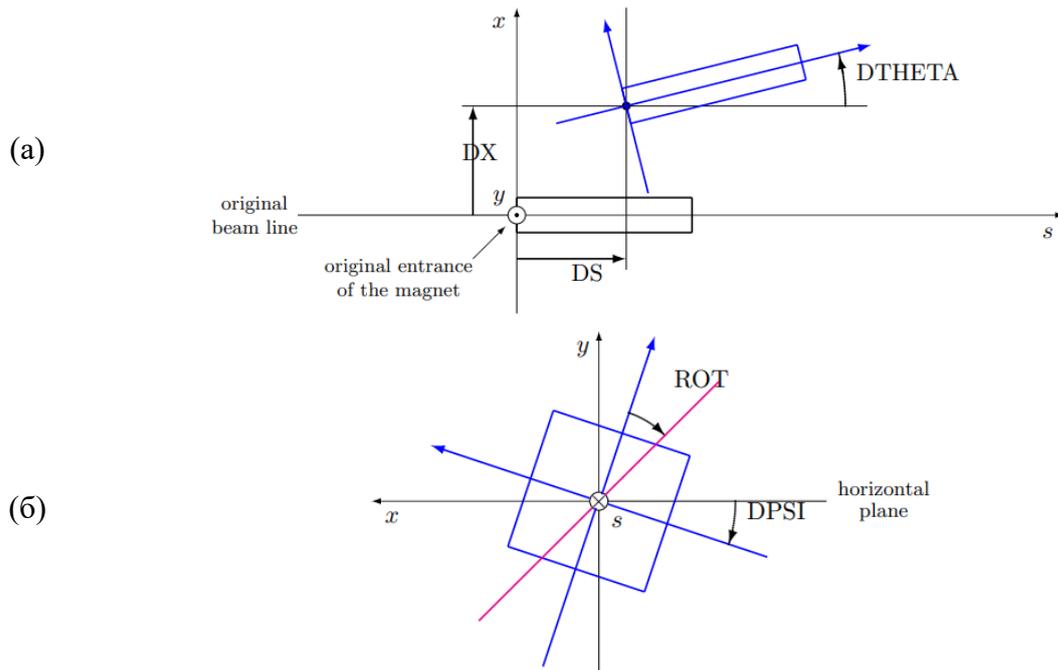


Рисунок 6. Параметры смещения магнитов относительно осей (x,s) – (а), (x,y) – (б)
[17]

Например, x -овые координаты луча в зависимости от смещения магнита вдоль оси x описываются следующими формулами [16]:

$$\begin{pmatrix} x_f \\ x'_f \end{pmatrix} = R \begin{pmatrix} x_i \\ x'_i \end{pmatrix} + [R - 1] \begin{pmatrix} DX \\ 0 \end{pmatrix}, \quad (4)$$

где R – передаточная матрица, x , x' – фазовые координаты, f – индекс, обозначающий конечную позицию. Первое слагаемое описывает положение пучка после прохождения магнита при отсутствии смещения. Однако к фазовым координатам пучка добавляется еще и дополнительное возмущение, описываемое формулой:

$$\vec{q} = [R - 1]\vec{m} = [R - 1] \begin{pmatrix} DX \\ 0 \end{pmatrix} \quad (5)$$

$\vec{m} = \begin{pmatrix} DX \\ 0 \end{pmatrix}$ – вектор несоосности [16].

В работе рассматриваются такие погрешности параметров магнитов как смещение магнитов в направлении x (DX), смещение магнитов в направлении y (DY), вращение дипольного магнита вокруг оси движения пучка частиц ($DPSI$).

Глава 3. Машинное обучение

Существует несколько определений понятия «машинное обучение». Артур Самуэль в 1959 году сформулировал, что машинное обучение (МО) – это научная дисциплина, которая наделяет компьютеры способностью учиться, не будучи явно запрограммированными [18]. Говоря простыми словами, это подход, предполагающий получение каких бы то ни было закономерностей из имеющихся данных. В последнее время к машинному обучению обращаются при наличии сложных задач, решение которых стандартными методами найти невозможно. Также МО используется в задачах с изменяющимися данными, т.к. система способна адаптироваться к новой вводимой информации и переучиваться на основе уже имеющихся закономерностей [19].

Машинное обучение имеет 3 составляющие:

- данные;
- признаки;
- алгоритм.

Данные – это любая предоставленная информация об объекте данных, на основе которой система будет обучаться. Наиболее удобным способом представления информации является таблица [20]. Исходные данные могут собираться как вручную, так и автоматически, однако их правильность и полнота в дальнейшем влияет на точность прогнозирования результатов.

Признаки – это наиболее важные характеристики, которые система должна отслуживать во время обучения. Выбор признаков является наиболее важным подготовительным этапом, именно здесь проводится исследовательский анализ, который помогает определить распределение данных, зависимость признаков друг от друга, наличие ошибок в данных. Для некоторых моделей машинного обучения важно, чтобы все значения в

признаках лежали в определенном числовом диапазоне, соответственно, необходимо привести их в нужный формат.

Алгоритм – выбор метода решения задачи. Именно от метода решения задачи зачатую зависит скорость, точность и правильность полученных результатов. Для оценки точности работы алгоритма принято устанавливать метрики, благодаря которым можно выбрать наиболее подходящую модель. Примеры: среднеквадратичное отклонение, средняя абсолютная ошибка и т.д. Выбор алгоритма подразумевает под собой и процесс оптимизации, который включает в себя перебор параметров модели, выбор подхода к решению задачи и т.д. [20]

Существуют различные виды машинного обучения в зависимости от постановки задачи:

- Обучение с учителем (supervised learning)
 - Классификация (classification)
 - Регрессия (regression)
 - Ранжирование (learning to rank)
- Обучение без учителя (unsupervised learning)
 - Кластеризация (clustering)
 - Уменьшение размерности (dimensionality reduction)
- Обучение с частичным привлечением учителя (semi-supervised learning)
- Обучение с подкреплением (reinforcement learning)

В работе применялось обучение с учителем, поэтому рассмотрим его более подробно.

При обучении с учителем задача ставится следующим образом: Каждый объект описывается парой x — данные (многомерный вектор), y — целевое значение, метка. Надо найти функцию $f(x) = y$.

Типичными задачами при обучении с учителем являются задачи классификации и задачи построения многомерной или одномерной регрессии. Первая описывается следующим образом: есть обучающая выборка, в которой представлены объекты в виде их признакового описания (вектор признаков) и метки класса. Надо найти такой алгоритм, который для каждого нового объекта определит, к какому классу он относится. Как правило, значения меток (классов) дискретны и принимают заранее определённые значения. Задача построения регрессии отличается от задачи классификации тем, что значения метки в ней непрерывно и находится внутри определенного диапазона. Необходимо найти такой алгоритм, что для каждого нового объекта (его признакового описания) он спрогнозирует значение целевой переменной. В работе рассматривается задача регрессии.

Как уже говорилось ранее оценка построенной модели обучения производится за счет метрик. Для регрессионной модели метрики, в основном описывают, насколько хорошо вписывается линия регрессии в данные. Наиболее часто используемые – это средний модуль ошибки, корень из среднеквадратичного значения ошибки [21].

Глава 4. Генерация данных и обучение моделей

4.1. Генерация данных и формирование датасета

Для формирования датасета необходимо сгенерировать данные, описывающие траекторию пучка и показатели различных ошибок на элементах кольца бустера. Генерация данных проводилась с помощью программы MAD-X.

MAD-X (Methodical Accelerator Design) – это проект с давней историей, нацеленный на задачи вычислительной физики в области проектирования и моделирования ускорителей частиц. Язык сценариев MAD-X сейчас является стандартом для моделирования и оптимизации динамики пучка заряженных частиц в ЦЕРНЕ. Код, содержащийся в файлах с расширением *.madx является C- подобным [22].

Генерация данных для каждой исследуемой погрешности магнита проводилась отдельно от других. Для этого был скорректирован исходный код, находящийся в файле main.madx. Данные генерировались только для дипольных магнитов и вычислялись исходя из Гауссова распределения, например, погрешность поворота магнита менялась в соответствии с формулой: $DPSI := 0,0001 * gauss()$. На выходе получали 1000 (или 5000) файлов с траекторией центра тяжести пучка (т.е. значениями x, y на пикап-мониторах) (рис.7) и столько же файлов с параметрами для элементов кольца (рис.8).

| * NAME | KEYWORD | X | Y |
|-----------|------------|------------------|------------------|
| \$ %s | %s | %le | %le |
| "BPMV1.1" | "VMONITOR" | -1.863866323e-09 | -5.435811784e-05 |
| "BPMH1.1" | "HMONITOR" | -2.199406682e-09 | -5.250097588e-05 |
| "BPMV2.1" | "VMONITOR" | -9.836743887e-10 | -6.665021719e-05 |
| "BPMH2.1" | "HMONITOR" | -1.003157042e-09 | -5.526403704e-05 |
| "BPMV3.1" | "VMONITOR" | 1.219676036e-09 | 9.253707165e-05 |
| "BPMH3.1" | "HMONITOR" | 1.422976514e-09 | 8.557438497e-05 |
| "BPMV4.1" | "VMONITOR" | -4.788042177e-10 | 0.0001216130124 |
| "BPMH4.1" | "HMONITOR" | -6.357792267e-10 | 0.0001060913201 |
| "BPMV5.1" | "VMONITOR" | -2.754344528e-09 | -2.030439934e-05 |
| "BPMH5.1" | "HMONITOR" | -3.122244786e-09 | -2.256090006e-05 |
| "BPMV6.1" | "VMONITOR" | -3.139227376e-09 | -0.0001336728706 |
| "BPMH6.1" | "HMONITOR" | -3.464080927e-09 | -0.0001194914179 |
| "BPMV1.2" | "VMONITOR" | -1.078839498e-10 | -5.885102503e-05 |
| "BPMH1.2" | "HMONITOR" | 1.778816512e-11 | -4.818237262e-05 |
| "BPMV2.2" | "VMONITOR" | 2.207088766e-09 | 9.909168664e-05 |
| "BPMH2.2" | "HMONITOR" | 2.504376015e-09 | 9.119297168e-05 |
| "BPMV3.2" | "VMONITOR" | -7.604703475e-10 | 0.0001177069733 |
| "BPMH3.2" | "HMONITOR" | -1.001901767e-09 | 0.0001023468483 |
| "BPMV4.2" | "VMONITOR" | -3.895274733e-09 | -2.917903132e-05 |
| "BPMH4.2" | "HMONITOR" | -4.40466903e-09 | -3.040353797e-05 |
| "BPMV5.2" | "VMONITOR" | -3.593507831e-09 | -0.0001350379428 |
| "BPMH5.2" | "HMONITOR" | -3.930024778e-09 | -0.0001204051018 |
| "BPMV6.2" | "VMONITOR" | 1.159626022e-10 | -5.102691609e-05 |

Рисунок 7. Файл с траекторией пучка

| @ NAME | %06s "EFIELD" | K0L | K0SL | K1L | K1SL | K2L | HETA | DPSI | MREX |
|------------------|---------------|-----|------|-----|------|-----|------|------------------|------|
| \$ %s | %s | %le | %le | %le | %le | %le | %le | %le | %le |
| "BOOSTER\$START" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QD1.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QD1.1" | | -0 | 0 | -0 | 0 | -0 | 0 | 0 | 0 |
| "QD1.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "BPMV1.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "ST1.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "BPMH1.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF1.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF1.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF1.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M2A.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | -8.057338117e-05 | 0 |
| "M2A.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M2A.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M2B.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | -2.428280337e-05 | 0 |
| "M2B.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M2B.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QD2.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QD2.1" | | -0 | 0 | -0 | 0 | -0 | 0 | 0 | 0 |
| "QD2.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "BPMV2.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "ST2.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "BPMH2.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF2.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF2.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QF2.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M3A.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M3A.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M3B.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M3B.1" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "M3B.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "QD3.1.ERR" | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Рисунок 8. Файл со значениями ошибок

Сокращения BPMH и BPMV (рис.7) используются для обозначения пикап-мониторов, считывающих вертикальную и горизонтальную составляющую движения пучка. На рис.8 можно увидеть поочередно все

элементы, имеющиеся в бустере. Для обозначения дипольных магнитов используется сокращение $MnA.k$ или $MnB.k$, где n – это номер регулярного периода, а k – номер суперпериода.

После генерации данных, проводится формирование датасета. Он состоит из параметра погрешности на магните и значения компоненты траектории. В ходе работы было выяснено, что параметр $DPSI$ является независимым от x -ого значения траектории, поэтому учитывать это в обучении бессмысленно. Зависимость одной величины от другой называется корреляцией и вычисляется по формуле:

$$cor = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (6)$$

где x_i, y_i – значения переменных, для которых определяется коэффициент корреляции, \bar{x}, \bar{y} – средние значения.

Корреляция x -вой компоненты траектории и DY равна 0.017, что говорит о несущественной взаимосвязи переменных, т.е. о том, что параметр DY почти никак не влияет на x -овую компоненту траектории. Для y -вой компоненты и погрешности смещения магнита вдоль оси y (DY) корреляция равна 0.974, а значит, что эти две компоненты сильно взаимосвязаны. Корреляция параметра Dx и x -овой компонентой траектории $cor(x, Dx) = -0.999$. Из этого следует обратно пропорциональная зависимость переменных. Влияние Dx на y -ковую компоненту траектории также можно не учитывать.

Таким образом, датасеты будут состоять из значений погрешностей параметров магнитов и значений компонент траектории, которые наилучшим образом коррелируют с погрешностями (рис.9)

| M2A.1_DX | M2B.1_DX | bpmh1_1 | bpmh2_1 | bpmh3_1 | bpmh4_1 | bpmh5_1 | bpmh6_1 | t |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---|
| -8,05734E-05 | -2,42828E-05 | 2,95582E-06 | 1,27301E-06 | -6,368E-06 | -5,04388E-06 | 3,38122E-06 | 7,04611E-06 | |
| -8,69784E-05 | -8,37268E-05 | 3,23493E-06 | 3,17235E-06 | -9,059E-06 | -8,53672E-06 | 4,00391E-06 | 1,09077E-05 | |
| 2,73418E-05 | 7,64226E-06 | -1,00257E-06 | -4,13286E-07 | 2,13821E-06 | 1,67945E-06 | -1,14371E-06 | -2,35671E-06 | |
| -0,000147781 | -5,41874E-05 | 5,4287E-06 | 2,63656E-06 | -1,20462E-05 | -9,76982E-06 | 6,26092E-06 | 1,34773E-05 | |
| -0,000100816 | 4,54089E-05 | 3,64023E-06 | -7,76789E-07 | -5,08869E-06 | -2,23653E-06 | 3,76431E-06 | 4,4656E-06 | |
| 0,000288505 | 6,04932E-05 | -1,05635E-05 | -3,73094E-06 | 2,17967E-05 | 1,66382E-05 | -1,19442E-05 | -2,3711E-05 | |
| 9,12398E-05 | 5,33113E-05 | -3,36694E-06 | -2,24856E-06 | 8,19164E-06 | 7,09934E-06 | -3,98768E-06 | -9,46068E-06 | |
| 2,29255E-05 | 7,80683E-05 | -8,95649E-07 | -2,58695E-06 | 4,51505E-06 | 5,26059E-06 | -1,39993E-06 | -6,08957E-06 | |
| -1,11427E-05 | -2,32411E-05 | 4,24031E-07 | 7,97679E-07 | -1,63595E-06 | -1,76643E-06 | 5,89946E-07 | 2,11577E-06 | |
| 8,65563E-05 | -6,30559E-05 | -3,10689E-06 | 1,41946E-06 | 3,45461E-06 | 6,26219E-07 | -3,08378E-06 | -2,45231E-06 | |
| -3,41899E-05 | -7,98616E-05 | 1,30764E-06 | 2,71487E-06 | -5,34446E-06 | -5,87964E-06 | 1,86277E-06 | 6,98271E-06 | |
| -0,000138973 | -7,30747E-05 | 5,12213E-06 | 3,1709E-06 | -1,21684E-05 | -1,03765E-05 | 6,02387E-06 | 1,39436E-05 | |
| 3,60214E-05 | -4,23561E-05 | -1,2806E-06 | 1,09454E-06 | 8,25517E-07 | -6,05703E-07 | -1,18419E-06 | -9,55276E-06 | |
| -0,000125371 | 4,49879E-06 | 4,56674E-06 | 6,58838E-07 | -8,3023E-06 | -5,57512E-06 | 5,00094E-06 | 8,53649E-06 | |
| -6,35956E-05 | 4,36891E-05 | 2,28474E-06 | -9,60375E-07 | -2,63848E-06 | -6,02026E-07 | 2,28198E-06 | 1,95333E-06 | |

Рисунок 9. Датасет для параметра DX двух магнитов для первых шести мониторов

4.2. Построение и обучение моделей

Первым шагом в процессе обучения модели является разделение данных на обучающее (train) и тестовое (test) множества. Обычно обучающее множество составляет 70% всего датасета. На нем происходит обучение модели, настройка численных параметров, оптимизация, выбор лучшей модели или изменение её структуры. В результате построения модели получаем веса, описывающие влияние каждого признака на выходной результат. В завершении обучения проверяем правильность и корректность модели на тестовом множестве, т.е. получаем оценку качества обученной модели.

Существуют различные обучающие модели: линейная регрессия, полиномиальная регрессия, многослойный перцептрон и др.

4.2.1. Линейная регрессионная модель

Линейная регрессионная модель – это одна из простейших моделей. Она пытается установить линейную связь между одной или несколькими переменными [21].

В общем виде линейная регрессия имеет вид:

$$\hat{y}_j = \theta_0 + \theta_1 * x_1 + \theta_2 * x_2 + \dots + \theta_n * x_n, \quad (7)$$

где θ_i – весовые коэффициенты модели, \hat{y}_j – j -ое спрогнозированное значение ($1 < j < k$), x_i – значение i -го признака, k – количество примеров в выборке, n – число признаков.

Признаками в данной работе являются значения компонент траектории, т.е. значения на пикап-мониторах.

В качестве метрик для описания качества модели были взяты значения RMSE – корень из среднеквадратичной ошибки и MAE – это среднее абсолютное значение ошибок. Здесь ошибки – это различия между предсказанными значениями (значениями, предсказанными нашей регрессионной моделью) и фактическими значениями переменной [23].

$$RMSE = \sqrt{\frac{1}{k} \sum_{j=1}^k (y_j - \hat{y}_j)^2} \quad (8)$$

$$MAE = \frac{1}{k} \sum_{j=1}^k |y_j - \hat{y}_j| \quad (9)$$

y_j – фактическое значение, \hat{y}_j – спрогнозированное значение, k – количество примеров в наборе данных, на которых измеряется ошибка.

Отличие RSME от MAE в том, что RSME дает представление о том, насколько большую ошибку система обычно допускает в своих прогнозах, с более высоким весом для крупных ошибок. Чем выше индекс нормы, тем больше она концентрируется на крупных значениях и пренебрегает мелкими значениями. Ошибка RMSE чувствительнее к выбросам, чем ошибка MAE.

4.2.2. Полиномиальная регрессионная модель

Полиномиальная регрессия используется, когда есть предположение, что имеющиеся данные составляют более сложный вид, чем прямая линия с некоторым шумом [21].

Полиномиальная регрессия описывается уравнением следующего вида:

$$\hat{y}_j = \theta_0 + \theta_1 * x_1^2 + \theta_2 * x_1 + \theta_3 * x_1 x_2 + \theta_4 * x_2 + \theta_5 * x_2^2 + \dots \quad (10)$$

По сравнению с уравнением линейной регрессии сюда добавляются квадратичные и перекрестные значения признаков. Метрики, используемые для оценки модели полиномиальной регрессии, аналогичны метрикам для линейной регрессии.

4.2.3. Многослойный перцептрон

Многослойный перцептрон – это один из видов искусственных нейронных сетей, а именно модель машинного обучения, в основе которой лежит устройство и работа нервной системы млекопитающих. Данная модель состоит из нескольких нейронов, связанных между собой и организованных в последовательные слои. Каждый слой возбуждается в зависимости от определенных условий.

Рассмотрим принцип работы обычного, однослойного перцептрона. Он состоит из входного вектора x , состоящего из n значений (признаков), и выходных значений 1, 0. Т.е. в процессе обучения необходимо определить функцию:

$$F(x) = \begin{cases} 1, & \text{если } wx + b > 0 \\ 0, & \text{в противном случае} \end{cases} \quad (11)$$

где w – вектор весов, wx – скалярное произведение $\sum_{i=1}^n w_i x_i$, b – смещение, F -функция, которую необходимо определить. Из математики известно, что выражение $wx + b = 0$ является выражением гиперплоскости, положение которой зависит от w и b . Если x лежит выше этой гиперплоскости, то 1, если ниже, то 0. Таким образом, процесс обучения сводится к поиску параметров w и b , т.е. имеет схожие части с линейной регрессией [23].

Математическую модель сети, которая состоит только из одного нейрона можно представить в виде двух последовательных операций:

1. Вычисление суммы входных сигналов с учетом их весов.
2. Применение функции активации к общей сумме, посчитанной ранее.

$$\hat{y}_j = \varphi \left(\sum_{i=1}^n \theta_i x_i^j + \theta_0 \right) \quad (12)$$

Здесь φ – функция активации, θ_i – весовые коэффициенты модели, \hat{y}_j – j -ое спрогнозированное значение ($1 < j < k$), x_i – значение i -го признака, k – количество примеров в выборке, n – число признаков.

Модель с несколькими линейными слоями называют многослойным перцептроном. Это сеть прямого распространения, состоящая из полносвязных слоев нейронов, как правило сигмоидальных (рис. 10). Под прямым распространением понимается, что сигнал идет от входа к выходу в одном направлении, не имея обратного хода. В полносвязном слое каждый нейрон связан со всеми нейронами из предыдущего слоя и из следующего. Сигмоидальный нейрон – это нейрон, имеющий сигмоидальную функцию активации [23].

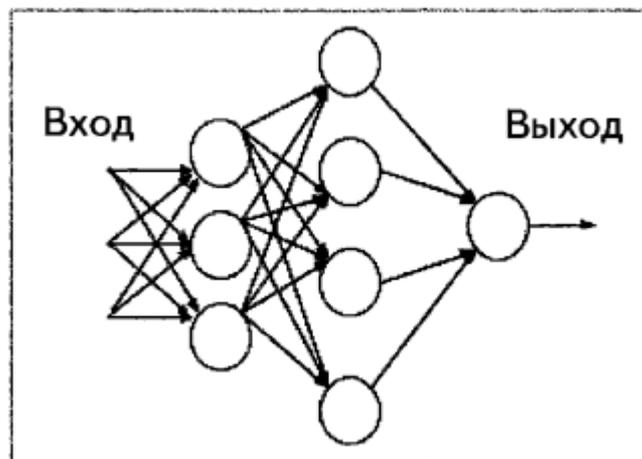


Рисунок 10. Графическое изображение модели многослового перцептрона с одним входным, одним промежуточным и одним выходным слоем [23].

Каждый нейрон, находящийся в первом слое, получает значение из входного вектора и возмущается в соответствии с определенными условиями, рассчитываются веса. Данные, полученные на выходе первого слоя, поступают на вход второго и т.д. Выходные данные последнего слоя являются выходными данными всей модели.

Функция активации применяется к выходу каждого нейрона для добавления нелинейности в вычисления, что позволяет обучать модель более сложным задачам. На рис. 11 представлена схема применения функции активации.

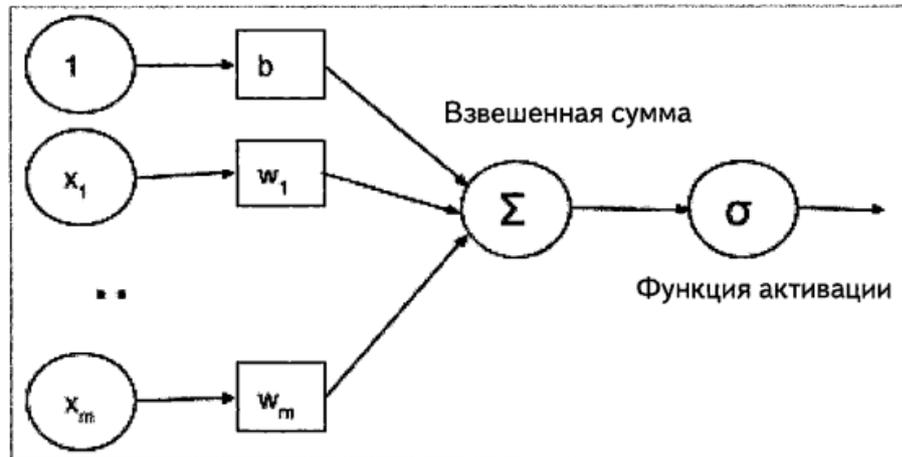


Рисунок 11. Схема применения функции активации

Функция определяет должен ли нейрон быть активирован или нет в зависимости от взвешенной суммы входных значений и порогового значения [24].

Существуют различные активизирующие функции, например, ступенчатая, линейная, функция ReLu, функция гиперболического тангенса и др. Но наиболее распространённой является сигмоидальная функция активации.

Она имеет следующий вид:

$$\varphi(x) = \frac{1}{1 + e^{-x}} \quad (13)$$

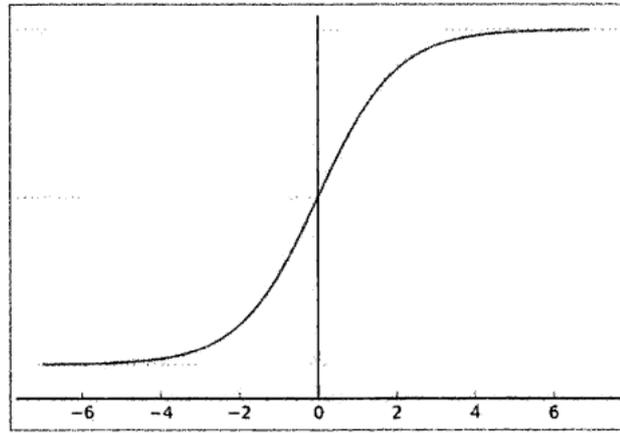


Рисунок 12. График сигмоидальной функции активации

Функция непрерывно изменяется от 0 до 1 при $x \in (-\infty; +\infty)$

Задача регрессии для многослойного перцептрона состоит в нахождении весов w при условии, что

1. $\mu(F(x, w), y_j) \leq \varepsilon_{train}$
2. $\mu(F(x, w), y_j) \leq \varepsilon_{test}$

Здесь x – входной вектор, $\hat{y}_j = F(x, w)$ – спрогнозированное значение, y_j – фактическое значение, $\varepsilon_{train}, \varepsilon_{test}$ – какие-то допустимые параметры, меньше которого должна быть метрика μ на обучающем/тестовом множествах.

Обучение многослойного перцептрона происходит следующим образом:

1. Выбор первоначальной архитектуры. Определение количества слоев, количества нейронов в одном слое (hidden), эпох обучения (epoch)[25].
2. Инициализация весов.
3. Выбор метода и параметров обучения. Наиболее часто встречающиеся параметры обучения – это пакетное (offline), когда на каждом шаге обучения на вход подаются все обучающие примеры; стохастическое (online) – на каждом шаге обучения, подается один произвольно выбранный обучающий набор; обучение в мини пакетах

(batch size) – на разных шагах подаются разные обучающие пакеты одинакового размера. Под методами подразумевается выбор оптимизатора и его параметров.

4. Обучение выбранной архитектуры.

5. Проверка по окончании обучения.

Машинное обучение проводилось с использованием библиотек Scikit-Learn, TensorFlow и Keras на языке программирования Python [26].

Глава 5. Результаты построения моделей

5.1. Вращение вокруг оси движения пучка частиц (DPSI)

При помощи программы для моделирования динамики пучка заряженных частиц MAD-X (Methodical Accelerator Design) [17,22] были сгенерированы данные, используемые для формирования датасета. В данном случае исследуем погрешность угла поворота магнита вокруг оси движения пучка частиц (*DPSI*).

$$DPSI = \begin{pmatrix} DPSI_1 \\ DPSI_2 \\ DPSI_3 \\ \vdots \\ DPSI_{40} \end{pmatrix},$$

где $DPSI_i$ – это погрешность угла поворота вокруг оси движения пучка для i -го магнита, все погрешности разные. Варьируя этот параметр, получаем различные значения координат пучка. Траектория движения фиксируется с помощью 48 пикап-мониторов. При формировании датасета выяснили, что значения на пикап-мониторах, считывающих горизонтальную составляющую движения пучка (BPMH), имеют практически нулевую корреляцию с *DPSI*, поэтому использовали значения только с вертикальных пикап мониторов (BPMV). Их 24 штуки во всем бустере.

$$BPMV = \begin{pmatrix} BPMV1_1 \\ BPMV2_1 \\ BPMV3_1 \\ \vdots \\ BPMV6_4 \end{pmatrix}$$

Построение модели обучения для 40 дипольных магнитов, имеющих в бустере, является не тривиальной задачей, тогда было принято решение о разделении задачи на несколько этапов, постепенно увеличивая количество магнитов, на которых меняется значение погрешности угла поворота. Таким образом, датасет состоит из вектора смещения угла поворота n -го количества магнитов и 24 y -координат с вертикальных пикап-мониторов. Длина датасета при одном/двух магнитах 1000 значений, при десяти и сорока магнитах 5000.

5.1.1. Один магнит

Первой построенной моделью была линейная регрессия. В качестве метрик для описания качества взяты значения RMSE и MAE, описанные ранее. Результаты построения линейной регрессии, представлены в таблице 3.

Таблица 3. Оценки качества модели линейной регрессии в случае изменения погрешности угла поворота (DPSI) для одного магнита

| | RMSE | | R ² | | MAE | |
|------------|-------------------|-------------------|----------------|------|-------------------|-------------------|
| | train | test | train | test | train | test |
| M2A.1_DPSI | $1,71 * 10^{-14}$ | $1,83 * 10^{-14}$ | 1,0 | 1,0 | $1,05 * 10^{-14}$ | $1,07 * 10^{-14}$ |

Построим график зависимости ошибок на обучающем и тестовом множествах от размера обучающей выборки.

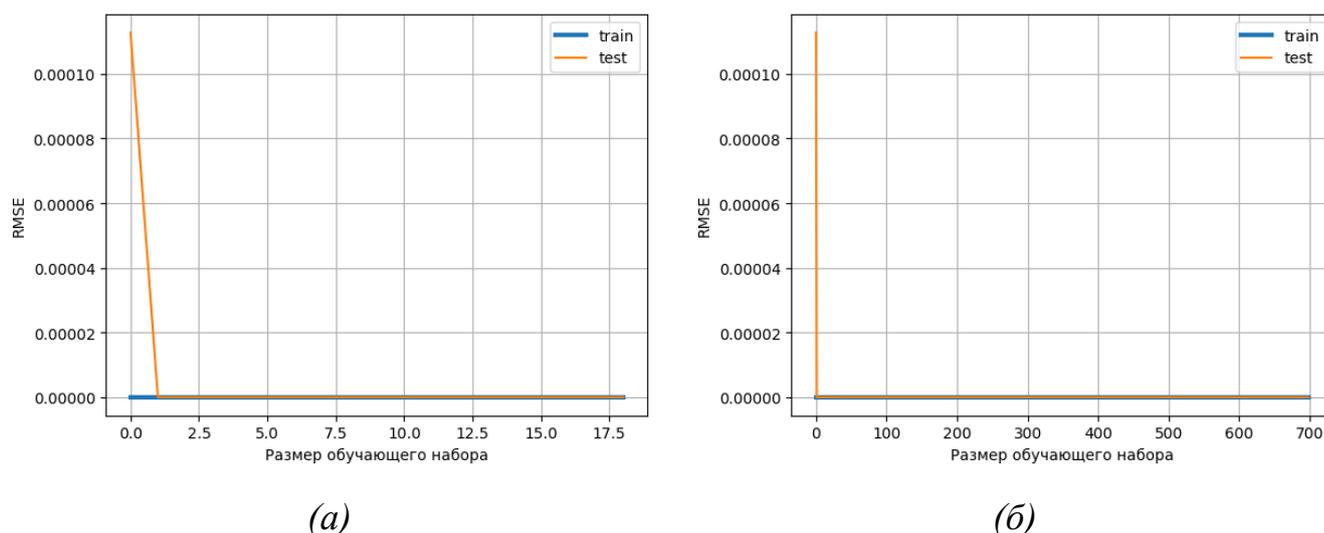


Рисунок 13. Кривые обучения для модели линейной регрессии в случае изменения погрешности угла поворота одного магнита (а) – обучающий набор с 20 примерами, (б)- полный обучающий набор (700 примеров)

Как видно из графиков (рис.13) в обучающем наборе почти все значения хорошо описываются моделью. И только одно значение из тестового множества не совсем корректно описывается выбранной моделью. Таким образом, можно сделать вывод о том, что параметр погрешности угла поворота вокруг оси движения пучка линейно зависит от параметров траектории.

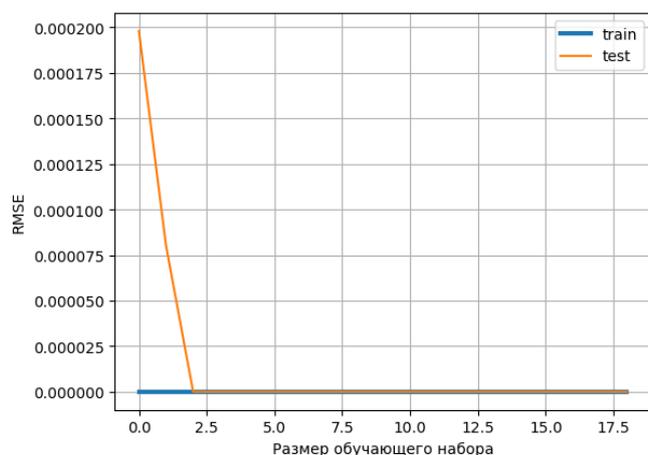
5.1.2. Два магнита

Следующим этапом стало построение обучающей модели, при изменении угла поворота у двух магнитов. В качестве второго магнита был взят следующий по порядку магнит в том же суперпериоде, что и первый. Мы придерживались тех же установок, что и в случае одного магнита.

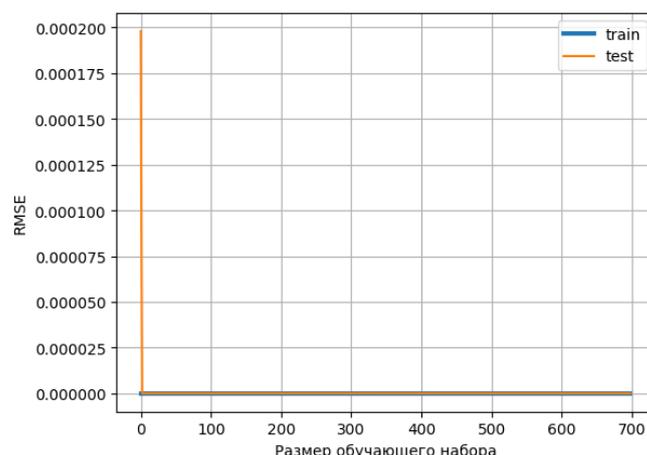
Построили линейную регрессию, получили следующие оценки качества.

Таблица 4. Оценки качества модели линейной регрессии в случае изменения погрешности угла поворота вокруг оси движения пучка (DPSI) для двух магнитов

| | RMSE | | MAE | |
|------------|-------------------|-------------------|-------------------|-------------------|
| | train | test | train | test |
| M2A.1_DPSI | $4,79 * 10^{-14}$ | $8,43 * 10^{-14}$ | $3,69 * 10^{-14}$ | $4,34 * 10^{-14}$ |
| M2B.1_DPSI | $3,81 * 10^{-14}$ | $5,74 * 10^{-14}$ | $2,65 * 10^{-14}$ | $2,88 * 10^{-14}$ |



(а)



(б)

Рисунок 14. Кривые обучения для модели линейной регрессии в случае изменения погрешности угла поворота двух магнитов (а) – обучающий набор с 20 примерами, (б) – полный обучающий набор (700 примеров)

Из рис. 14 и таблицы 4 видно, что модель линейной регрессии очень хорошо описывает имеющиеся данные, т.е. предсказанные такой моделью значения будут иметь высокую точность. Так же построили полиномиальную регрессию степени 2. Сделали это с помощью встроенного класса

PolynomialFeatures. Полученные ошибки на тестовом и обучающем множествах представлены в таблице 5.

Таблица 5. Оценки качества модели полиномиальной регрессии в случае изменения погрешности угла поворота вокруг оси движения пучка (DPSI) для двух магнитов

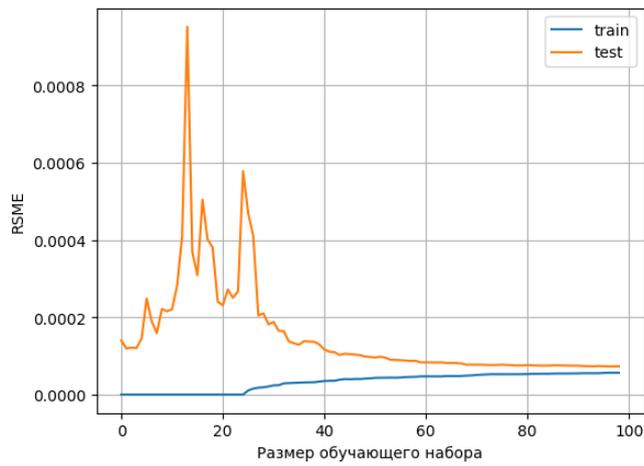
| | RMSE | | MAE | |
|------------|-------------------|-------------------|-------------------|-------------------|
| | train | test | train | test |
| M2A.1_DPSI | $1,57 * 10^{-14}$ | $7,47 * 10^{-14}$ | $9,99 * 10^{-15}$ | $1,63 * 10^{-14}$ |
| M2B.1_DPSI | $1,45 * 10^{-14}$ | $5,54 * 10^{-14}$ | $9,32 * 10^{-15}$ | $1,64 * 10^{-14}$ |

Как видно в таблице 5 значения ошибок для модели полиномиальной регрессии меньше, чем значения, полученные при построении линейной регрессии. Следовательно, полиномиальная регрессия является лучшим обобщением для имеющихся данных.

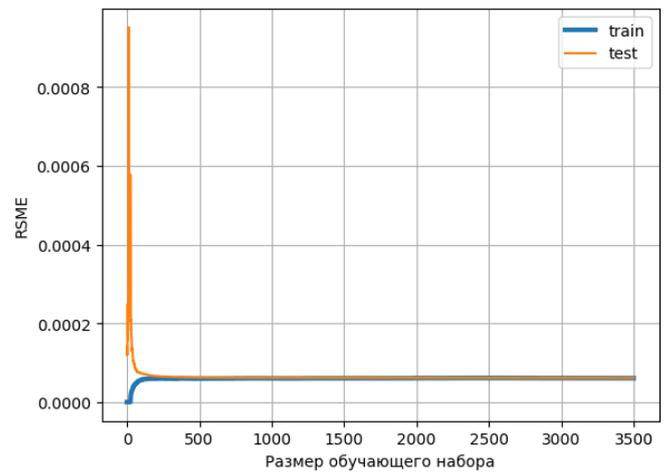
5.1.3. Десять магнитов

В дальнейшем построили модели линейной регрессии и многослойного перцептрона в случае, когда погрешность угла поворота вокруг оси движения пучка варьируется для 10 магнитов. При этом величина датасета была взята равная 5000. Следовательно, число обучающих примеров увеличилось.

Построим график обучающей кривой. Данный график представляет собой зависимость производительности модели на обучающем и тестовом множествах от размера обучающего набора. Для этого несколько раз была построена модель с разным размером обучающего множества. Как видно по кривой для обучающего набора, в нем присутствуют около 20 образцов, к которым модель в полной мере может быть подогнана (рис.15, а), что и объясняет начало кривой с ошибкой равной нулю. По мере добавления примеров в обучающий набор идеальное описание становится невозможно из-за нелинейности остальных данных и увеличения их зашумленности, значение ошибки увеличивается, пока не стабилизируется (рис.15, б).



(а)



(б)

Рисунок 15. Кривые обучения для модели линейной регрессии в случае изменения погрешности угла поворота десяти магнитов (а) – обучающий набор со 100 примерами, (б)- полный обучающий набор (3500 примеров)

По производительности модели на тестовом множестве можно сказать следующее: пока модель обучалась на небольшом количестве данных ошибка была большая и не постоянная, что говорит о плохой обобщающей способности модели. С увеличением обучающих данных, обещающая способность модели растет, ошибка на тестовом множестве уменьшается и приближается к ошибке на обучающем множестве.

Таблица 6. Оценки качества модели линейной регрессии в случае изменения погрешности угла поворота вокруг оси движения пучка (DPSI) для десяти магнитов

| | RMSE | | MAE | |
|------------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DPSI | $2,59 * 10^{-5}$ | $2,56 * 10^{-5}$ | $2,08 * 10^{-5}$ | $2,06 * 10^{-5}$ |
| M2B.1_DPSI | $5,10 * 10^{-5}$ | $5,18 * 10^{-5}$ | $4,15 * 10^{-5}$ | $4,15 * 10^{-5}$ |
| M3A.1_DPSI | $7,76 * 10^{-5}$ | $7,86 * 10^{-5}$ | $6,23 * 10^{-5}$ | $6,29 * 10^{-5}$ |
| M3B.1_DPSI | $5,42 * 10^{-5}$ | $5,54 * 10^{-5}$ | $4,33 * 10^{-5}$ | $4,42 * 10^{-5}$ |
| M4A.1_DPSI | $8,21 * 10^{-5}$ | $8,55 * 10^{-5}$ | $6,54 * 10^{-5}$ | $6,83 * 10^{-5}$ |
| M4B.1_DPSI | $5,38 * 10^{-5}$ | $5,53 * 10^{-5}$ | $4,31 * 10^{-5}$ | $4,43 * 10^{-5}$ |
| M5A.1_DPSI | $8,10 * 10^{-5}$ | $8,23 * 10^{-5}$ | $6,36 * 10^{-5}$ | $6,56 * 10^{-5}$ |
| M5B.1_DPSI | $5,23 * 10^{-5}$ | $5,26 * 10^{-5}$ | $4,17 * 10^{-5}$ | $4,20 * 10^{-5}$ |

| | | | | |
|------------|------------------|------------------|------------------|------------------|
| M6A.1_DPSI | $8,31 * 10^{-5}$ | $8,30 * 10^{-5}$ | $6,55 * 10^{-5}$ | $6,63 * 10^{-5}$ |
| M6B.1_DPSI | $4,14 * 10^{-5}$ | $4,19 * 10^{-5}$ | $3,27 * 10^{-5}$ | $3,31 * 10^{-5}$ |

Анализируя значения ошибок из таблицы 6, можно сказать, что значения ошибок имеют порядок величин близкий к погрешностям угла поворота вокруг оси движения пучка будут и прогнозироваться с точностью менее 80%.

Рассмотрим еще одну модель – многослойный перцептрон. Используем следующие параметры модели обучения: 3 слоя + один выходной, количество нейронов в скрытых слоях 160, 2 в выходном слое, количество эпох обучения 200, функция активации softsign, которая имеет следующий вид:

$$\varphi(x) = \frac{x}{|x| + 1} \quad (14)$$

Построенная модель многослойного перцептрона имела значение ошибки RMSE больше, чем значения погрешности DPSI в датасете. Из-за этого было принято решение оптимизировать модель методом стохастического градиентного спуска. Таким образом, в обучающую модель был введен параметр optimizer = adam [27]. Новая модель аналогичным образом не дала хороший результат, поэтому было увеличено количество скрытых слоев до 4, и количество нейронов в одном слое до 200. Полученные значения ошибок представлены в таблице 7.

На рис.16 для модели многослойного перцептрона показан график зависимости ошибки RSME от эпохи обучения для тестовой и обучающей выборок, содержащих значения погрешности угла поворота десяти магнитов. Исходя из графика, можно сказать, что ошибка спадает с увеличением эпох достаточно быстро.

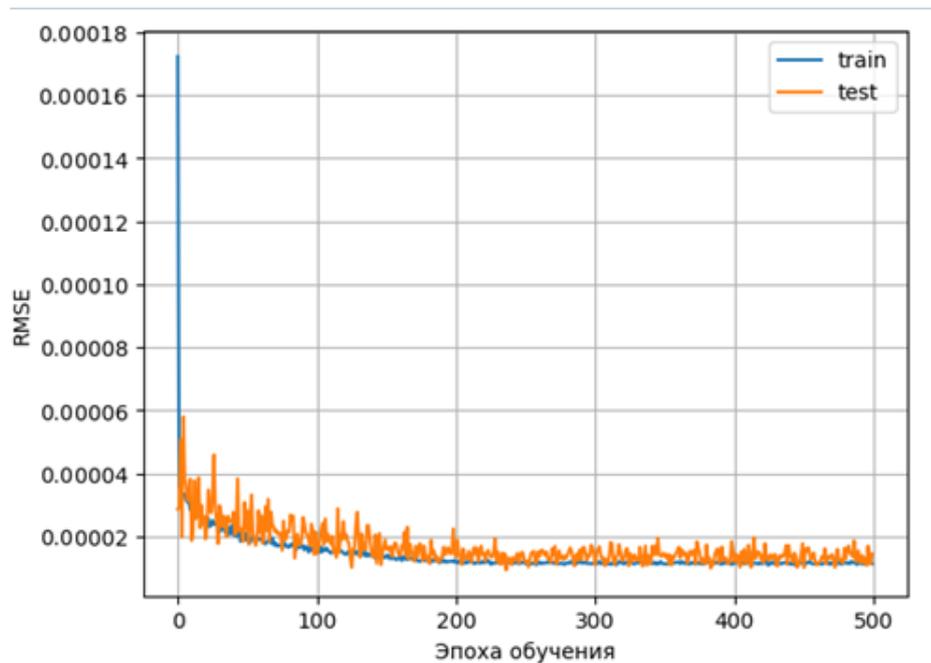


Рисунок 16. График зависимости ошибки RMSE от эпохи обучения модели многослойного перцептрона для тестовой и обучающей выборок, содержащих значения погрешности угла поворота десяти магнитов

Таблица 7. Оценки качества модели многослойного перцептрона в случае изменения погрешности угла поворота вокруг оси движения пучка (DPSI) для десяти магнитов

| | RMSE | | MAE | |
|------------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DPSI | $1,19 * 10^{-5}$ | $1,99 * 10^{-6}$ | $9,67 * 10^{-6}$ | $9,87 * 10^{-6}$ |
| M2B.1_DPSI | $9,14 * 10^{-6}$ | $9,30 * 10^{-6}$ | $7,33 * 10^{-6}$ | $7,47 * 10^{-6}$ |
| M3A.1_DPSI | $1,97 * 10^{-5}$ | $1,93 * 10^{-5}$ | $1,77 * 10^{-5}$ | $1,72 * 10^{-5}$ |
| M3B.1_DPSI | $1,37 * 10^{-5}$ | $1,38 * 10^{-5}$ | $1,16 * 10^{-5}$ | $1,16 * 10^{-5}$ |
| M4A.1_DPSI | $1,20 * 10^{-5}$ | $1,22 * 10^{-5}$ | $9,83 * 10^{-6}$ | $9,88 * 10^{-6}$ |
| M4B.1_DPSI | $2,41 * 10^{-5}$ | $2,45 * 10^{-5}$ | $2,25 * 10^{-5}$ | $2,29 * 10^{-5}$ |
| M5A.1_DPSI | $1,44 * 10^{-5}$ | $1,44 * 10^{-5}$ | $1,22 * 10^{-5}$ | $1,22 * 10^{-5}$ |
| M5B.1_DPSI | $8,79 * 10^{-6}$ | $9,07 * 10^{-6}$ | $6,69 * 10^{-6}$ | $7,25 * 10^{-6}$ |
| M6A.1_DPSI | $1,50 * 10^{-5}$ | $1,57 * 10^{-5}$ | $1,27 * 10^{-5}$ | $1,34 * 10^{-5}$ |
| M6B.1_DPSI | $1,00 * 10^{-5}$ | $1,00 * 10^{-5}$ | $7,99 * 10^{-6}$ | $8,08 * 10^{-6}$ |

Исходя из значений в таблицах 6 и 7, можно сказать, что модель линейной регрессии и модель многослойного перцептрона имеют один порядок оценки качества.

5.1.4. Сорок магнитов

Для всех магнитов мы использовали ту же модель многослойного перцептрона, что и ранее. График зависимости ошибки MAE от эпохи обучения, представлен на рис.17. Исходя из этого графика можно сказать, что среднее абсолютное значение ошибок имеет порядок около $1,00 * 10^{-4}$, что не попадает в допустимый уровень погрешности, равный 20%, а значит, что модель линейной регрессии описывает данные для 40 магнитов чуть лучше.

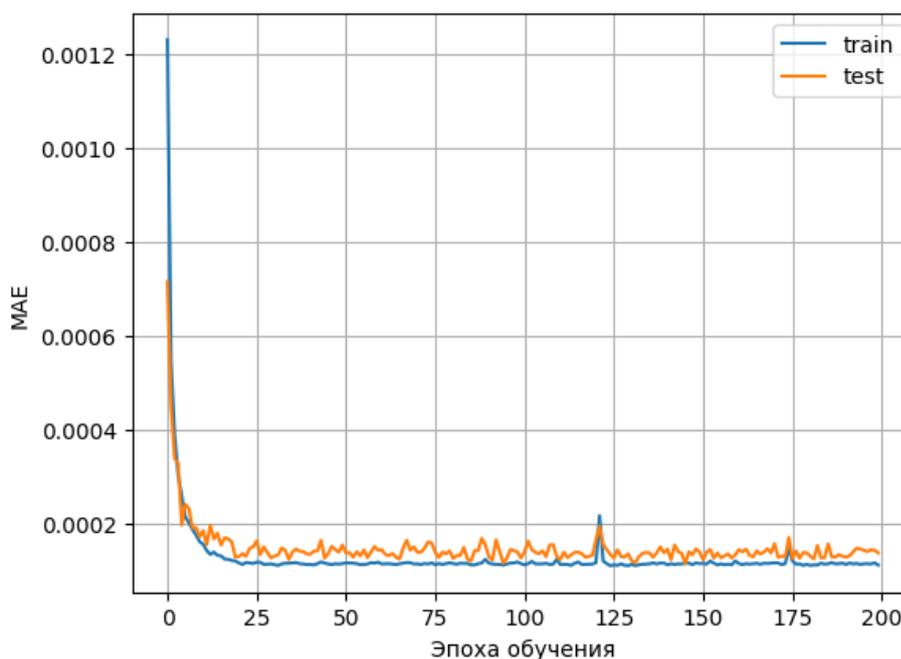


Рисунок 17. График зависимости ошибки MAE от эпохи обучения для тестовой и обучающей выборок, содержащих значения погрешности угла поворота сорока магнитов

5.2. Смещение магнита в направлении x (DX)

Для погрешности смещения магнита (DX) построили такие же модели, как и для погрешности угла поворота вокруг оси x (DPSI). Как упоминалось в главе про формирование датасета корреляция данного параметра и у-ковой составляющей траектории близка к нулю, поэтому ее учитывать не имеет смысла. Таким образом, датасет состоит из значений погрешностей на n магнитах и двадцати четырех значений на горизонтальных пикап-мониторах.

5.2.1. Два магнита

Для начала провели обучение линейной регрессии, получили следующие ошибки на обучающем и тестовом множестве (таблица 8).

Таблица 8. Оценки качества модели линейной регрессии в случае смещения магнита вдоль оси x для двух магнитов

| | RMSE | | R ² | | MAE | |
|----------|-------------------|-------------------|----------------|------|-------------------|-------------------|
| | train | test | train | test | train | test |
| M2A.1_DX | $4,51 * 10^{-13}$ | $5,82 * 10^{-13}$ | 1,0 | 1,0 | $2,07 * 10^{-13}$ | $2,66 * 10^{-13}$ |
| M2B.1_DX | $2,93 * 10^{-13}$ | $3,74 * 10^{-13}$ | 1,0 | 1,0 | $1,52 * 10^{-13}$ | $1,93 * 10^{-13}$ |

Определили параметры модели, т.е. весовые коэффициенты модели:

Таблица 9. Весовые коэффициенты модели линейной регрессии в случае смещения двух магнитов вдоль оси x.

| Коэффициент | M2A.1_DX | M2B.1_DX |
|---------------|-------------------|--------------------|
| θ_0 | $1,65 * 10^{-13}$ | $-1,51 * 10^{-13}$ |
| θ_1 | 140,22 | -69,89 |
| θ_2 | 11,50 | -17,11 |
| θ_3 | -8,56 | 8,06 |
| θ_4 | -13,42 | 12,29 |
| θ_5 | 1,76 | -11,77 |
| θ_6 | 2,11 | -1,87 |
| θ_7 | -7,89 | -10,79 |
| θ_8 | -38,48 | 38,05 |
| θ_9 | -17,79 | 19,00 |
| θ_{10} | 33,70 | 18,63 |
| θ_{11} | 0,12 | 0,73 |
| θ_{12} | 69,46 | -48,63 |
| θ_{13} | 10,86 | -1,61 |
| θ_{14} | -2,51 | 5,30 |
| θ_{15} | -59,77 | 11,32 |
| θ_{16} | -20,48 | 16,30 |
| θ_{17} | -258,90 | 289,53 |
| θ_{18} | -1,07 | 5,29 |

| | | |
|---------------|---------|--------|
| θ_{19} | -22,83 | 24,36 |
| θ_{20} | -235,02 | 213,25 |
| θ_{21} | -19,77 | 14,43 |
| θ_{22} | 68,91 | -70,38 |
| θ_{23} | 14,30 | -3,40 |
| θ_{24} | -26,00 | 23,70 |

Из значений ошибок на обучающем и тестовом множестве (таблица 8) видно, что значения x -вой компоненты траектории движения пучка заряженных частиц линейно зависит от погрешности смещения магнита в направлении оси x .

В дальнейшем была построена и модель полиномиальной регрессии степени 2. Значения корня из среднеквадратичного отклонения, среднего абсолютного значения ошибок и параметр R^2 представлены в таблице 10.

Таблица 10. Оценки качества модели полиномиальной регрессии в случае изменения погрешности смещения магнита вдоль оси x для двух магнитов

| | RMSE | | R^2 | | MAE | |
|----------|-------------------|-------------------|-------|------|-------------------|-------------------|
| | train | test | train | test | train | test |
| M2A.1_DX | $4,20 * 10^{-13}$ | $5,71 * 10^{-13}$ | 1,0 | 1,0 | $1,62 * 10^{-13}$ | $2,27 * 10^{-13}$ |
| M2B.1_DX | $2,63 * 10^{-13}$ | $3,54 * 10^{-13}$ | 1,0 | 1,0 | $1,00 * 10^{-13}$ | $1,39 * 10^{-13}$ |

По значениям ошибок можно сказать, что полиномиальная регрессия описывается представленные данные лучше, чем линейная регрессия.

5.2.2. Десять магнитов

При изменении погрешности смещения магнита вдоль оси x на 10 магнитах аналогично предыдущим пунктам построили линейную и полиномиальную модель. Значения метрик на тестовом и обучающих множествах показаны в таблицах ниже (таблица 11,12).

Построили график обучающей кривой для модели линейной регрессии. Данный график представляет собой зависимость производительности модели на обучающем и тестовом множествах от размера обучающего набора. Для этого несколько раз была построена модель с разным размером обучающего множества. Как видно по кривой для обучающего набора, в нем присутствуют около 20 образцов, к которым модель в полной мере может быть подогнана (рис.18, а), что и объясняет начало кривой с ошибкой равной нулю. По мере добавления данных идеальная подгонка данных становится невозможна из-за нелинейности остальных данных и увеличения их зашумленности, значение ошибки ползет вверх, пока не стабилизируется (рис.18, б). По производительности модели на тестовом множестве можно сказать следующее: пока модель обучалась на небольшом количестве данных ошибка была большая и не постоянная. Поскольку модель видит все больше обучающих данных, ошибка на тестовом множестве уменьшается и приближается к ошибке на обучающем множестве.

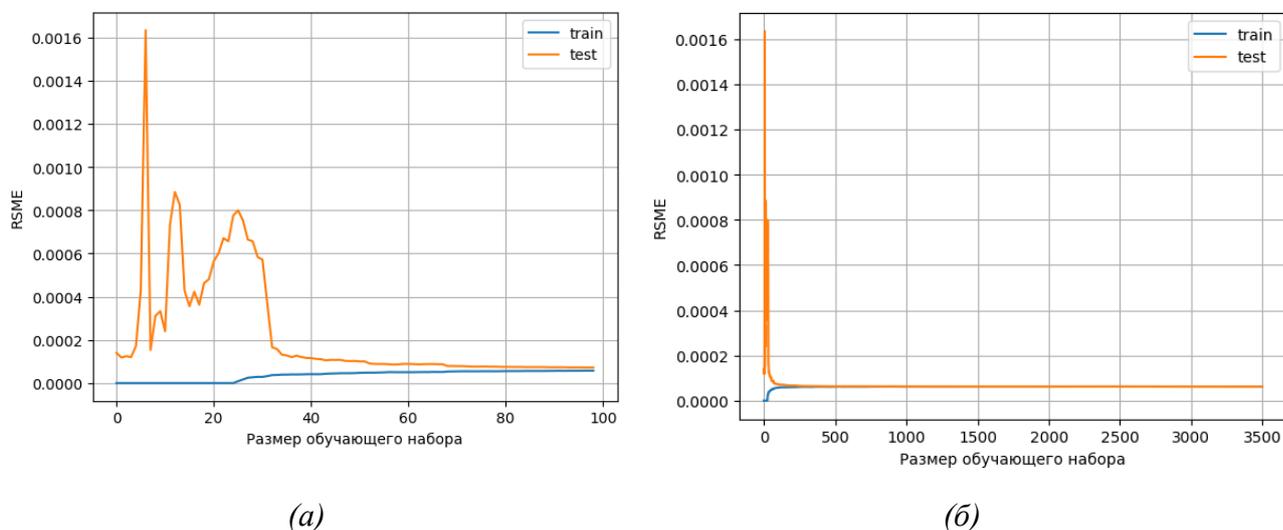


Рисунок 18. Кривые обучения модели линейной регрессии в случае изменения погрешности смещения вдоль оси x десяти магнитов. (а) – в приближении, где количество обучающих примеров меньше 100, (б) – общие графики кривых обучения

Таблица 11. Оценки качества модели линейной регрессии в случае изменения погрешности смещения магнита вдоль оси x для десяти магнитов

| | RMSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DX | $5,22 * 10^{-5}$ | $6,73 * 10^{-5}$ | $4,54 * 10^{-5}$ | $5,62 * 10^{-5}$ |
| M2B.1_DX | $5,17 * 10^{-5}$ | $8,64 * 10^{-5}$ | $4,15 * 10^{-5}$ | $7,24 * 10^{-5}$ |
| M3A.1_DX | $5,53 * 10^{-5}$ | $9,55 * 10^{-5}$ | $8,26 * 10^{-5}$ | $8,65 * 10^{-5}$ |
| M3B.1_DX | $5,35 * 10^{-5}$ | $4,0 * 10^{-5}$ | $3,87 * 10^{-5}$ | $3,27 * 10^{-5}$ |
| M4A.1_DX | $8,14 * 10^{-5}$ | $8,34 * 10^{-5}$ | $6,68 * 10^{-5}$ | $4,18 * 10^{-5}$ |
| M4B.1_DX | $1,28 * 10^{-4}$ | $5,25 * 10^{-5}$ | $9,33 * 10^{-5}$ | $4,35 * 10^{-5}$ |
| M5A.1_DX | $7,76 * 10^{-5}$ | $9,46 * 10^{-5}$ | $6,84 * 10^{-5}$ | $8,34 * 10^{-5}$ |
| M5B.1_DX | $8,48 * 10^{-5}$ | $6,35 * 10^{-5}$ | $5,76 * 10^{-5}$ | $5,85 * 10^{-5}$ |
| M6A.1_DX | $7,95 * 10^{-5}$ | $5,63 * 10^{-5}$ | $5,77 * 10^{-5}$ | $4,86 * 10^{-5}$ |
| M6B.1_DX | $4,43 * 10^{-5}$ | $8,71 * 10^{-5}$ | $3,23 * 10^{-5}$ | $7,08 * 10^{-5}$ |

Таблица 12. Оценки качества модели полиномиальной регрессии в случае изменения погрешности смещения магнита вдоль оси x для десяти магнитов

| | RMSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DX | $5,75 * 10^{-5}$ | $6,15 * 10^{-5}$ | $5,31 * 10^{-5}$ | $5,29 * 10^{-5}$ |
| M2B.1_DX | $5,36 * 10^{-5}$ | $5,43 * 10^{-5}$ | $4,34 * 10^{-5}$ | $7,46 * 10^{-5}$ |
| M3A.1_DX | $8,47 * 10^{-6}$ | $3,25 * 10^{-5}$ | $7,15 * 10^{-5}$ | $2,95 * 10^{-5}$ |
| M3B.1_DX | $5,45 * 10^{-5}$ | $4,36 * 10^{-5}$ | $3,90 * 10^{-5}$ | $3,23 * 10^{-5}$ |
| M4A.1_DX | $5,17 * 10^{-5}$ | $5,33 * 10^{-5}$ | $4,69 * 10^{-5}$ | $4,45 * 10^{-5}$ |
| M4B.1_DX | $5,12 * 10^{-5}$ | $5,48 * 10^{-5}$ | $9,68 * 10^{-5}$ | $4,73 * 10^{-5}$ |
| M5A.1_DX | $6,05 * 10^{-5}$ | $6,04 * 10^{-5}$ | $7,17 * 10^{-5}$ | $9,18 * 10^{-5}$ |
| M5B.1_DX | $6,87 * 10^{-5}$ | $6,34 * 10^{-5}$ | $4,86 * 10^{-5}$ | $5,86 * 10^{-5}$ |
| M6A.1_DX | $4,39 * 10^{-5}$ | $5,46 * 10^{-5}$ | $6,04 * 10^{-5}$ | $4,74 * 10^{-5}$ |
| M6B.1_DX | $4,46 * 10^{-5}$ | $7,87 * 10^{-5}$ | $3,21 * 10^{-5}$ | $7,72 * 10^{-5}$ |

Исходя из результатов в таблицах 11,12 видно, что значения ошибок на обучающих и тестовых множествах примерно равны среднему значению смещения магнита вдоль оси x, а значит возможно существование модели, лучше описывающей имеющиеся данные.

Затем была построена модель многослойного перцептрона. В обучении было проведено 200 эпох, на каждой из которых модель видит по 16 обучающих примера прежде, чем обновить веса. Число нейронов в одном слое равно 100, всего 3 скрытых слоя. Функция активации – softsign (14). Функция потерь MSE. Также проводилось построение модели многослойного перцептрона со всеми вышеуказанными параметрами и с метрикой MAE. Оценки качества моделей для каждого магнита можно увидеть в таблице 13.

Таблица 13. Оценки качества моделей многослойного перцептрона с метриками MSE и MAE в случае изменения погрешности смещения магнита вдоль оси x для десяти магнитов

| | MSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | Test |
| M2A.1_DX | $2,38 * 10^{-9}$ | $2,38 * 10^{-9}$ | $2,23 * 10^{-5}$ | $2,56 * 10^{-5}$ |
| M2B.1_DX | $1,62 * 10^{-9}$ | $1,55 * 10^{-9}$ | $1,35 * 10^{-5}$ | $1,37 * 10^{-5}$ |
| M3A.1_DX | $2,39 * 10^{-9}$ | $2,71 * 10^{-9}$ | $2,02 * 10^{-5}$ | $2,43 * 10^{-5}$ |
| M3B.1_DX | $1,89 * 10^{-9}$ | $1,89 * 10^{-9}$ | $1,46 * 10^{-5}$ | $1,77 * 10^{-5}$ |
| M4A.1_DX | $1,67 * 10^{-9}$ | $1,63 * 10^{-9}$ | $1,42 * 10^{-5}$ | $1,18 * 10^{-5}$ |
| M4B.1_DX | $2,78 * 10^{-9}$ | $2,76 * 10^{-9}$ | $2,30 * 10^{-5}$ | $2,40 * 10^{-5}$ |
| M5A.1_DX | $1,92 * 10^{-9}$ | $2,17 * 10^{-9}$ | $1,53 * 10^{-5}$ | $1,81 * 10^{-5}$ |
| M5B.1_DX | $1,89 * 10^{-9}$ | $1,89 * 10^{-9}$ | $1,70 * 10^{-5}$ | $1,78 * 10^{-5}$ |
| M6A.1_DX | $2,04 * 10^{-9}$ | $2,71 * 10^{-9}$ | $1,77 * 10^{-5}$ | $1,43 * 10^{-5}$ |
| M6B.1_DX | $1,76 * 10^{-9}$ | $2,37 * 10^{-9}$ | $1,50 * 10^{-5}$ | $1,96 * 10^{-5}$ |

В таблице 13 видно, что ошибка MAE имеет примерно тот же порядок величины, что и значение смещения магнита относительно оси x. Для получения параметра RMSE из оценки MSE необходимо извлечь квадратный

корень. Соответственно получаем, что ошибки тестового и обучающего множества имеют порядок 10^{-5} .

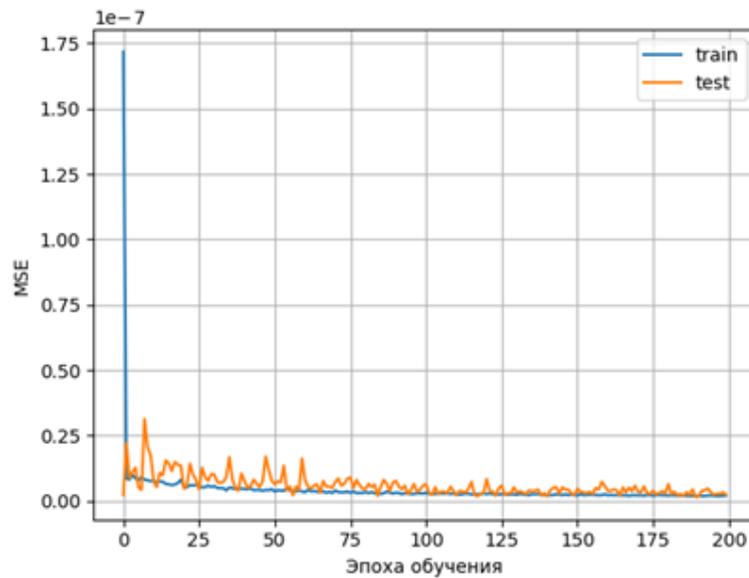


Рисунок 19. График зависимости среднеквадратичной ошибки от эпохи обучения модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси x для десяти магнитов

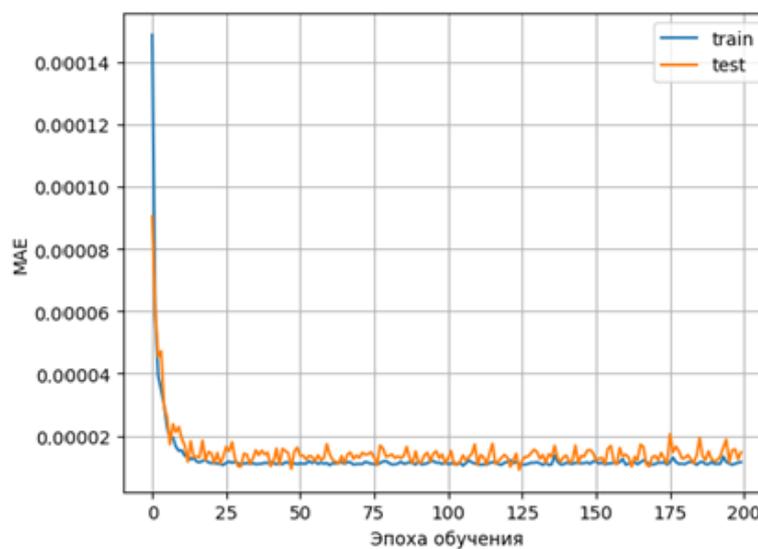


Рисунок 20. График зависимости средней абсолютной ошибки от эпохи обучения модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси x для десяти магнитов

Как видно на рис.19, на первых эпохах обучения ошибка на обучающем множестве значительно превосходит как ошибку на тестовом множестве, так и ошибки других эпох обучения. Это может говорить о том, что начальные веса были подобраны некорректно. Далее погрешность обучения спадает и

стабилизируется. Для тестового множества разброс погрешностей модели на первых эпохах значительно больше, чем на более поздних. Это может говорить о слабости модели в начале обучения.

На рис. 20 для обучающего множества видно, что ошибка спадает с увеличением эпох обучения. Для тестового множества видим колеблющееся значения метрики, в среднем, с большим значениями ошибки, чем у обучающего набора.

5.2.3. Сорок магнитов

По аналогии с предыдущими подпунктами для 40 магнитов были построены 2 вида моделей: линейная регрессия, модель многослойного перцептрона. Данные по линейной регрессии можно увидеть ниже.

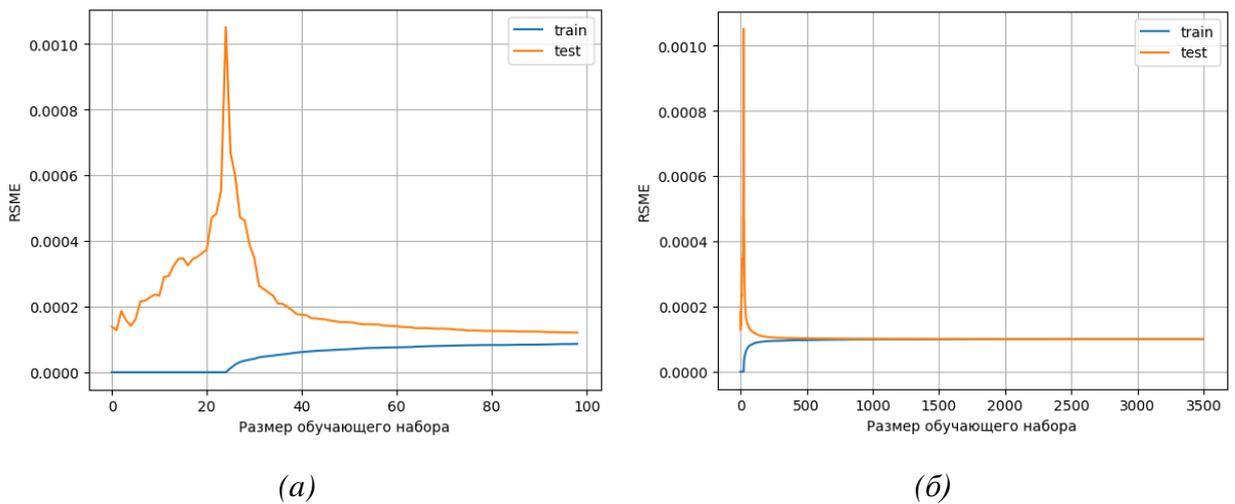


Рисунок 21. Кривые обучения линейной регрессии в случае изменения погрешности смещения вдоль оси x для 40 магнитов (а) – в приближении, где количество обучающих примеров меньше 100, (б) – общие графики кривых обучения

Таблица 14. Оценки качества модели линейной регрессии в случае изменения погрешности смещения магнита вдоль оси x для сорока магнитов

| | MSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DX | $9,80 * 10^{-9}$ | $9,61 * 10^{-9}$ | $7,77 * 10^{-5}$ | $7,88 * 10^{-5}$ |
| M2B.1_DX | $1,02 * 10^{-8}$ | $1,00 * 10^{-8}$ | $8,01 * 10^{-5}$ | $8,07 * 10^{-5}$ |
| M3A.1_DX | $1,02 * 10^{-8}$ | $9,70 * 10^{-9}$ | $7,83 * 10^{-5}$ | $8,09 * 10^{-5}$ |

| | | | | |
|----------|------------------|------------------|------------------|------------------|
| M3B.1_DX | $9,81 * 10^{-9}$ | $9,47 * 10^{-9}$ | $7,88 * 10^{-5}$ | $7,74 * 10^{-5}$ |
| M4A.1_DX | $9,89 * 10^{-9}$ | $1,09 * 10^{-8}$ | $7,96 * 10^{-5}$ | $8,34 * 10^{-5}$ |
| M4B.1_DX | $1,00 * 10^{-8}$ | $9,98 * 10^{-9}$ | $7,98 * 10^{-5}$ | $7,95 * 10^{-5}$ |
| M5A.1_DX | $9,84 * 10^{-9}$ | $9,59 * 10^{-9}$ | $7,93 * 10^{-5}$ | $7,95 * 10^{-5}$ |
| M5B.1_DX | $1,01 * 10^{-8}$ | $9,26 * 10^{-9}$ | $8,08 * 10^{-5}$ | $7,67 * 10^{-5}$ |
| M6A.1_DX | $1,01 * 10^{-8}$ | $1,03 * 10^{-8}$ | $7,96 * 10^{-5}$ | $8,05 * 10^{-5}$ |
| M6B.1_DX | $9,93 * 10^{-9}$ | $1,03 * 10^{-8}$ | $7,96 * 10^{-5}$ | $8,19 * 10^{-5}$ |
| M2A.2_DX | $1,02 * 10^{-8}$ | $9,80 * 10^{-9}$ | $8,09 * 10^{-5}$ | $7,97 * 10^{-5}$ |
| M2B.2_DX | $1,01 * 10^{-8}$ | $9,80 * 10^{-9}$ | $7,95 * 10^{-5}$ | $7,97 * 10^{-5}$ |
| M3A.2_DX | $9,96 * 10^{-9}$ | $1,01 * 10^{-8}$ | $7,99 * 10^{-5}$ | $7,90 * 10^{-5}$ |
| M3B.2_DX | $1,02 * 10^{-8}$ | $1,01 * 10^{-8}$ | $8,05 * 10^{-5}$ | $8,05 * 10^{-5}$ |
| M4A.2_DX | $9,96 * 10^{-9}$ | $9,79 * 10^{-9}$ | $7,87 * 10^{-5}$ | $7,95 * 10^{-5}$ |
| M4B.2_DX | $9,93 * 10^{-9}$ | $1,04 * 10^{-8}$ | $7,97 * 10^{-5}$ | $8,13 * 10^{-5}$ |
| M5A.2_DX | $1,00 * 10^{-8}$ | $1,03 * 10^{-8}$ | $8,02 * 10^{-5}$ | $8,18 * 10^{-5}$ |
| M5B.2_DX | $1,00 * 10^{-8}$ | $9,71 * 10^{-9}$ | $7,97 * 10^{-5}$ | $7,97 * 10^{-5}$ |
| M6A.2_DX | $9,72 * 10^{-9}$ | $1,11 * 10^{-8}$ | $7,80 * 10^{-5}$ | $8,46 * 10^{-5}$ |
| M6B.2_DX | $9,97 * 10^{-9}$ | $1,01 * 10^{-8}$ | $7,95 * 10^{-5}$ | $8,00 * 10^{-5}$ |
| M2A.3_DX | $1,01 * 10^{-8}$ | $9,63 * 10^{-9}$ | $8,02 * 10^{-5}$ | $7,82 * 10^{-5}$ |
| M2B.3_DX | $9,95 * 10^{-9}$ | $9,26 * 10^{-9}$ | $7,95 * 10^{-5}$ | $7,92 * 10^{-5}$ |
| M3A.3_DX | $1,01 * 10^{-8}$ | $1,06 * 10^{-8}$ | $8,03 * 10^{-5}$ | $8,28 * 10^{-5}$ |
| M3B.3_DX | $1,02 * 10^{-8}$ | $1,01 * 10^{-8}$ | $8,03 * 10^{-5}$ | $8,07 * 10^{-5}$ |
| M4A.3_DX | $9,93 * 10^{-9}$ | $1,04 * 10^{-8}$ | $7,94 * 10^{-5}$ | $8,19 * 10^{-5}$ |
| M4B.3_DX | $9,49 * 10^{-9}$ | $9,97 * 10^{-9}$ | $7,74 * 10^{-5}$ | $7,97 * 10^{-5}$ |
| M5A.3_DX | $9,95 * 10^{-9}$ | $9,90 * 10^{-9}$ | $7,96 * 10^{-5}$ | $7,97 * 10^{-5}$ |
| M5B.3_DX | $9,78 * 10^{-9}$ | $1,07 * 10^{-8}$ | $7,88 * 10^{-5}$ | $8,27 * 10^{-5}$ |
| M6A.3_DX | $1,04 * 10^{-8}$ | $9,48 * 10^{-9}$ | $8,11 * 10^{-5}$ | $7,93 * 10^{-5}$ |
| M6B.3_DX | $9,83 * 10^{-9}$ | $9,97 * 10^{-9}$ | $7,96 * 10^{-5}$ | $8,04 * 10^{-5}$ |
| M2A.4_DX | $1,01 * 10^{-8}$ | $1,02 * 10^{-8}$ | $8,00 * 10^{-5}$ | $8,07 * 10^{-5}$ |
| M2B.4_DX | $9,75 * 10^{-9}$ | $9,78 * 10^{-9}$ | $7,85 * 10^{-5}$ | $7,91 * 10^{-5}$ |
| M3A.4_DX | $9,71 * 10^{-9}$ | $1,02 * 10^{-8}$ | $7,82 * 10^{-5}$ | $8,14 * 10^{-5}$ |
| M3B.4_DX | $1,01 * 10^{-8}$ | $9,91 * 10^{-9}$ | $8,05 * 10^{-5}$ | $7,87 * 10^{-5}$ |
| M4A.4_DX | $1,02 * 10^{-8}$ | $9,93 * 10^{-9}$ | $8,06 * 10^{-5}$ | $8,07 * 10^{-5}$ |

| | | | | |
|----------|------------------|------------------|------------------|------------------|
| M4B.4_DX | $1,02 * 10^{-8}$ | $9,86 * 10^{-9}$ | $7,91 * 10^{-5}$ | $8,06 * 10^{-5}$ |
| M5A.4_DX | $9,87 * 10^{-9}$ | $1,04 * 10^{-8}$ | $7,92 * 10^{-5}$ | $8,10 * 10^{-5}$ |
| M5B.4_DX | $1,01 * 10^{-8}$ | $1,04 * 10^{-8}$ | $7,99 * 10^{-5}$ | $8,15 * 10^{-5}$ |
| M6A.4_DX | $9,59 * 10^{-9}$ | $1,00 * 10^{-8}$ | $7,78 * 10^{-5}$ | $8,02 * 10^{-5}$ |
| M6B.4_DX | $9,92 * 10^{-9}$ | $9,90 * 10^{-9}$ | $7,98 * 10^{-5}$ | $8,00 * 10^{-5}$ |

Как видно из графика кривых обучения в множестве train присутствует около 20 образцов, к которым модель в полной мере может быть подогнана (рис.21, а), что и объясняет начало кривой с ошибкой равной нулю. По мере добавления данных идеальная подгонка становится невозможна из-за нелинейности остальных данных и увеличения их зашумленности, значение ошибки увеличивается, пока не стабилизируется (рис.21, б). По производительности модели на тестовом множестве можно сказать следующее: при малом количестве обучающих примеров, модель плохо прогнозирует тестовые значения. При добавлении нелинейных данных в обучающую выборку значение ошибки RMSE плавно, но быстро стремится к значениям на обучающем множестве.

Построим модель многослойного перцептрона со следующими параметрами: 4 слоя + 1 выходной, количество нейронов в каждом слое 160, функция активации softsign (формула 14). Оптимизатор модели – adam. В качестве метрик были взяты MSE и MAE. Графики изменения ошибок в зависимости от эпохи обучения представлены на рис.22, рис.23.

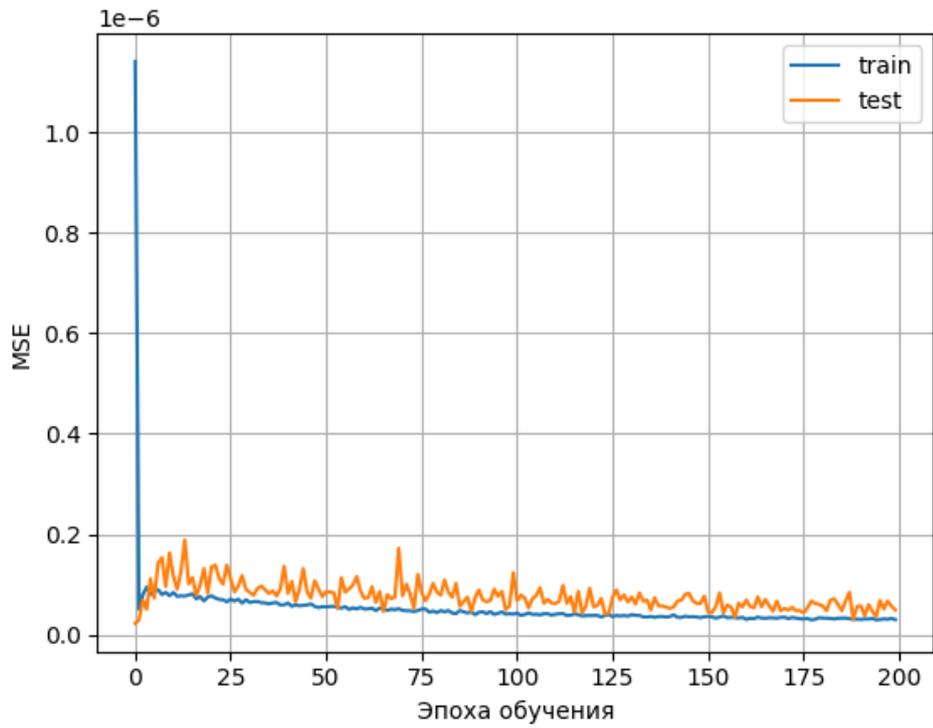


Рисунок 22. График зависимости среднеквадратичной ошибки от эпохи обучения модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси x для сорока магнитов

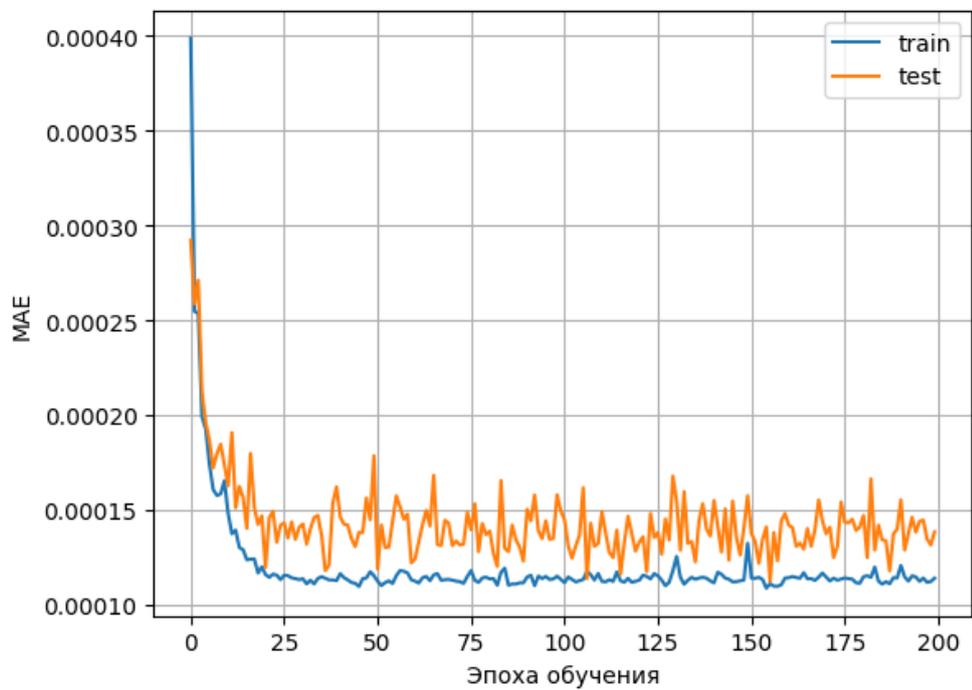


Рисунок 23. Зависимости ошибки MAE от эпохи обучения для модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси x для 40 магнитов

Таблица 15. Оценки качества модели многослойного перцептрона с метриками MSE и MAE в случае изменения погрешности смещения магнита вдоль оси x для сорока магнитов

| | <i>MSE</i> | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DX | $1,02 * 10^{-7}$ | $1,03 * 10^{-7}$ | $3,05 * 10^{-4}$ | $3,07 * 10^{-4}$ |
| M2B.1_DX | $3,60 * 10^{-8}$ | $3,65 * 10^{-8}$ | $1,65 * 10^{-4}$ | $1,65 * 10^{-4}$ |
| M3A.1_DX | $2,08 * 10^{-8}$ | $1,89 * 10^{-8}$ | $1,18 * 10^{-4}$ | $1,18 * 10^{-4}$ |
| M3B.1_DX | $6,06 * 10^{-8}$ | $6,00 * 10^{-8}$ | $2,26 * 10^{-4}$ | $2,25 * 10^{-4}$ |
| M4A.1_DX | $1,95 * 10^{-8}$ | $1,99 * 10^{-8}$ | $1,15 * 10^{-4}$ | $1,17 * 10^{-4}$ |
| M4B.1_DX | $1,37 * 10^{-8}$ | $1,40 * 10^{-8}$ | $9,42 * 10^{-5}$ | $9,25 * 10^{-5}$ |
| M5A.1_DX | $3,65 * 10^{-8}$ | $3,60 * 10^{-8}$ | $1,67 * 10^{-4}$ | $1,66 * 10^{-4}$ |
| M5B.1_DX | $1,12 * 10^{-8}$ | $1,11 * 10^{-8}$ | $7,98 * 10^{-5}$ | $8,49 * 10^{-5}$ |
| M6A.1_DX | $7,78 * 10^{-8}$ | $7,86 * 10^{-8}$ | $2,60 * 10^{-4}$ | $2,62 * 10^{-4}$ |
| M6B.1_DX | $6,69 * 10^{-8}$ | $6,73 * 10^{-8}$ | $2,39 * 10^{-4}$ | $2,40 * 10^{-4}$ |
| M2A.2_DX | $1,03 * 10^{-8}$ | $9,66 * 10^{-9}$ | $8,12 * 10^{-5}$ | $7,90 * 10^{-5}$ |
| M2B.2_DX | $2,14 * 10^{-8}$ | $2,16 * 10^{-8}$ | $1,21 * 10^{-4}$ | $1,29 * 10^{-4}$ |
| M3A.2_DX | $9,63 * 10^{-8}$ | $9,63 * 10^{-8}$ | $2,94 * 10^{-4}$ | $2,99 * 10^{-4}$ |
| M3B.2_DX | $4,42 * 10^{-8}$ | $4,43 * 10^{-8}$ | $1,87 * 10^{-4}$ | $1,85 * 10^{-4}$ |
| M4A.2_DX | $5,03 * 10^{-8}$ | $5,12 * 10^{-8}$ | $2,03 * 10^{-4}$ | $2,05 * 10^{-4}$ |
| M4B.2_DX | $5,14 * 10^{-8}$ | $4,95 * 10^{-8}$ | $2,05 * 10^{-4}$ | $1,99 * 10^{-4}$ |
| M5A.2_DX | $2,46 * 10^{-8}$ | $2,43 * 10^{-8}$ | $1,32 * 10^{-4}$ | $1,33 * 10^{-4}$ |
| M5B.2_DX | $1,69 * 10^{-8}$ | $1,71 * 10^{-8}$ | $1,03 * 10^{-4}$ | $9,97 * 10^{-5}$ |
| M6A.2_DX | $1,23 * 10^{-8}$ | $1,35 * 10^{-8}$ | $8,84 * 10^{-5}$ | $9,32 * 10^{-5}$ |
| M6B.2_DX | $4,84 * 10^{-8}$ | $5,15 * 10^{-8}$ | $1,98 * 10^{-4}$ | $2,05 * 10^{-4}$ |
| M2A.3_DX | $2,27 * 10^{-8}$ | $2,30 * 10^{-8}$ | $9,02 * 10^{-5}$ | $9,06 * 10^{-5}$ |
| M2B.3_DX | $1,55 * 10^{-8}$ | $1,59 * 10^{-8}$ | $9,85 * 10^{-5}$ | $1,01 * 10^{-4}$ |
| M3A.3_DX | $2,51 * 10^{-8}$ | $2,52 * 10^{-8}$ | $1,32 * 10^{-4}$ | $1,33 * 10^{-4}$ |
| M3B.3_DX | $1,03 * 10^{-8}$ | $1,00 * 10^{-8}$ | $8,03 * 10^{-5}$ | $8,09 * 10^{-5}$ |
| M4A.3_DX | $1,94 * 10^{-7}$ | $1,93 * 10^{-8}$ | $4,30 * 10^{-4}$ | $4,37 * 10^{-4}$ |
| M4B.3_DX | $1,14 * 10^{-8}$ | $1,17 * 10^{-8}$ | $8,53 * 10^{-5}$ | $8,76 * 10^{-5}$ |
| M5A.3_DX | $1,32 * 10^{-8}$ | $1,36 * 10^{-8}$ | $9,18 * 10^{-5}$ | $9,14 * 10^{-5}$ |
| M5B.3_DX | $1,89 * 10^{-7}$ | $1,87 * 10^{-7}$ | $4,24 * 10^{-4}$ | $4,20 * 10^{-4}$ |

| | | | | |
|----------|------------------|------------------|------------------|------------------|
| M6A.3_DX | $1,57 * 10^{-8}$ | $1,55 * 10^{-8}$ | $9,66 * 10^{-5}$ | $1,00 * 10^{-4}$ |
| M6B.3_DX | $8,87 * 10^{-8}$ | $8,89 * 10^{-8}$ | $2,81 * 10^{-4}$ | $2,88 * 10^{-4}$ |
| M2A.4_DX | $2,19 * 10^{-8}$ | $2,20 * 10^{-8}$ | $1,23 * 10^{-4}$ | $1,23 * 10^{-4}$ |
| M2B.4_DX | $9,90 * 10^{-9}$ | $9,99 * 10^{-9}$ | $7,91 * 10^{-5}$ | $7,93 * 10^{-5}$ |
| M3A.4_DX | $3,25 * 10^{-8}$ | $3,27 * 10^{-8}$ | $1,56 * 10^{-4}$ | $1,56 * 10^{-4}$ |
| M3B.4_DX | $1,00 * 10^{-7}$ | $1,03 * 10^{-7}$ | $3,00 * 10^{-4}$ | $3,06 * 10^{-4}$ |
| M4A.4_DX | $9,14 * 10^{-8}$ | $8,97 * 10^{-8}$ | $2,85 * 10^{-4}$ | $2,85 * 10^{-4}$ |
| M4B.4_DX | $1,63 * 10^{-8}$ | $1,65 * 10^{-8}$ | $1,04 * 10^{-4}$ | $1,01 * 10^{-4}$ |
| M5A.4_DX | $1,26 * 10^{-8}$ | $1,25 * 10^{-7}$ | $3,40 * 10^{-4}$ | $3,48 * 10^{-4}$ |
| M5B.4_DX | $5,39 * 10^{-8}$ | $5,36 * 10^{-8}$ | $2,11 * 10^{-4}$ | $2,17 * 10^{-4}$ |
| M6A.4_DX | $1,06 * 10^{-8}$ | $1,06 * 10^{-7}$ | $3,10 * 10^{-4}$ | $3,09 * 10^{-4}$ |
| M6B.4_DX | $2,41 * 10^{-8}$ | $2,43 * 10^{-8}$ | $1,30 * 10^{-4}$ | $1,32 * 10^{-4}$ |

Как видно из значений в таблице 15 ошибка на тестовом и обучающем множествах при построении модели многослойного перцептрона для 40 магнитов имеет примерно такой же порядок, что и сами значения предсказываемого параметра, а соответственно полученное значение может не попадать в допустимый предел погрешности 20%.

Модель линейной регрессии чуть лучше описывает наши данные, чем модель многослойного перцептрона.

5.3. Смещение магнита в направлении y (DY)

Для погрешности смещения магнита (DY) построили такие же модели, как и для погрешности угла поворота вокруг оси x (DPSI). Как упоминалось в главе про формирование датасета корреляция данного параметра и y -ковой составляющей траектории близка к нулю, поэтому её учитывать не имеет смысла. Таким образом, датасет состоит из значений погрешностей на n магнитах и двадцати четырех значений на вертикальных пикап-мониторах.

5.3.1. Два магнита

Для начала провели обучение линейной регрессии, получили следующие ошибки на обучающем и тестовом множестве (таблица 16).

Таблица 16. Оценки качества модели линейной регрессии в случае смещения магнита вдоль оси у для двух магнитов

| | RMSE | | R ² | | MAE | |
|----------|-------------------|-------------------|----------------|------|-------------------|-------------------|
| | train | test | train | test | train | test |
| M2A.1_DY | $1,84 * 10^{-13}$ | $1,88 * 10^{-13}$ | 1,0 | 1,0 | $1,24 * 10^{-13}$ | $1,25 * 10^{-13}$ |
| M2B.1_DY | $2,18 * 10^{-13}$ | $2,25 * 10^{-13}$ | 1,0 | 1,0 | $1,31 * 10^{-13}$ | $1,39 * 10^{-13}$ |

Из значений ошибок на обучающем и тестовом множестве (таблица 16) видно, что значения у-вой компоненты траектории движения пучка заряженных частиц линейно зависит от погрешности смещения магнита в направлении оси у.

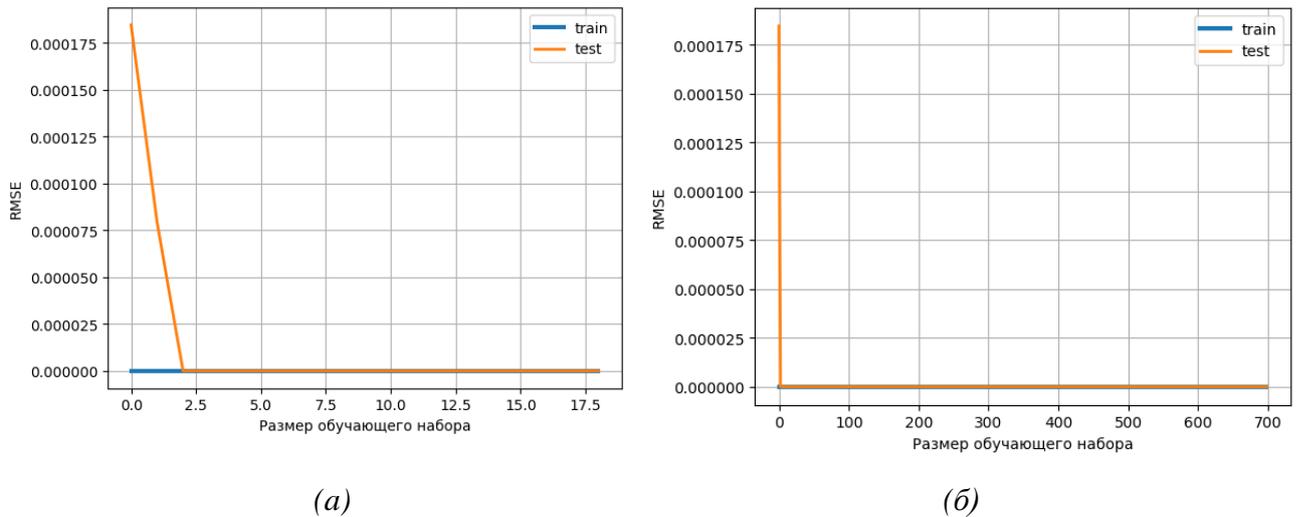


Рисунок 24. Кривые обучения линейной регрессии в случае изменения погрешности смещения магнита вдоль оси у для двух магнитов (а) – в приближении, где количество обучающих примеров меньше 20, (б) – общие графики кривых обучения

Как видно из графиков на рис.24 линейная регрессия полностью описывает обучающее множество, о чем говорит близкое к нулю значение ошибки RMSE. Для тестового множества ошибка велика на первых двух примерах, после чего стабилизируется.

В дальнейшем была построена и модель полиномиальной регрессии степени 2. Значения корня из среднеквадратичного отклонения, среднего абсолютного значения ошибок и параметр R^2 представлены в таблице 17.

Таблица 17. Оценки качества модели полиномиальной регрессии в случае смещения магнита вдоль оси y для двух магнитов

| | RMSE | | R^2 | | MAE | |
|----------|-------------------|-------------------|-------|------|-------------------|-------------------|
| | train | test | train | test | train | test |
| M2A.1_DY | $8,28 * 10^{-14}$ | $8,63 * 10^{-14}$ | 1,0 | 1,0 | $5,60 * 10^{-14}$ | $5,80 * 10^{-14}$ |
| M2B.1_DY | $8,08 * 10^{-14}$ | $8,45 * 10^{-14}$ | 1,0 | 1,0 | $5,34 * 10^{-14}$ | $5,71 * 10^{-14}$ |

По значениям ошибок можно сказать, что полиномиальная регрессия описывается представленные данные лучше, чем линейная регрессии, т.к. значение оценки качества меньше.

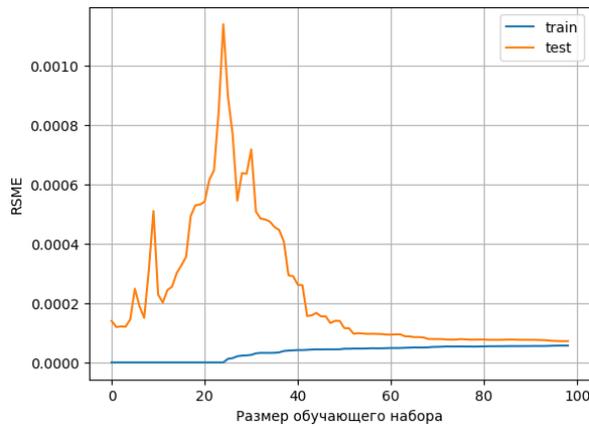
5.3.2. Десять магнитов

При изменении погрешности смещения магнита вдоль оси y на 10 магнитах аналогично предыдущим пунктам построили линейную и полиномиальную. Значения ошибок на тестовом и обучающих множествах для модели линейной регрессии показаны в таблице 18.

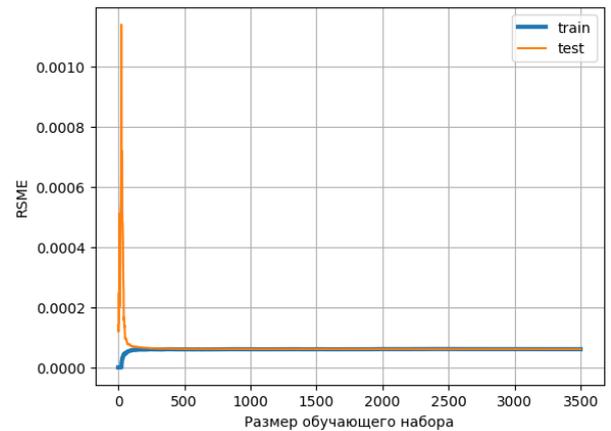
Таблица 18. Оценки качества модели линейной регрессии в случае изменения погрешности смещения магнита вдоль оси y для десяти магнитов

| | RMSE | | R^2 | | MAE | |
|----------|------------------|------------------|-------|------|------------------|------------------|
| | train | test | train | test | train | test |
| M2A.1_DY | $2,60 * 10^{-5}$ | $2,67 * 10^{-5}$ | 0,93 | 0,93 | $2,08 * 10^{-5}$ | $2,07 * 10^{-5}$ |
| M2B.1_DY | $5,17 * 10^{-5}$ | $5,17 * 10^{-5}$ | 0,73 | 0,73 | $4,13 * 10^{-5}$ | $4,15 * 10^{-5}$ |
| M3A.1_DY | $7,80 * 10^{-5}$ | $7,86 * 10^{-5}$ | 0,40 | 0,37 | $6,30 * 10^{-5}$ | $6,37 * 10^{-5}$ |
| M3B.1_DY | $5,43 * 10^{-5}$ | $5,54 * 10^{-5}$ | 0,70 | 0,67 | $4,33 * 10^{-5}$ | $4,42 * 10^{-5}$ |
| M4A.1_DY | $8,23 * 10^{-5}$ | $8,55 * 10^{-5}$ | 0,32 | 0,32 | $6,55 * 10^{-5}$ | $6,83 * 10^{-5}$ |
| M4B.1_DY | $5,39 * 10^{-4}$ | $5,53 * 10^{-5}$ | 0,71 | 0,69 | $4,31 * 10^{-5}$ | $4,44 * 10^{-5}$ |
| M5A.1_DY | $8,23 * 10^{-5}$ | $8,21 * 10^{-5}$ | 0,32 | 0,31 | $6,56 * 10^{-5}$ | $6,56 * 10^{-5}$ |
| M5B.1_DY | $5,27 * 10^{-5}$ | $5,24 * 10^{-5}$ | 0,73 | 0,70 | $4,20 * 10^{-5}$ | $4,28 * 10^{-5}$ |
| M6A.1_DY | $8,32 * 10^{-5}$ | $8,31 * 10^{-5}$ | 0,32 | 0,32 | $6,64 * 10^{-5}$ | $6,67 * 10^{-5}$ |

| | | | | | | |
|----------|------------------|------------------|------|------|------------------|------------------|
| M6B.1_DY | $4,15 * 10^{-5}$ | $4,14 * 10^{-5}$ | 0,82 | 0,83 | $3,31 * 10^{-5}$ | $3,38 * 10^{-5}$ |
|----------|------------------|------------------|------|------|------------------|------------------|



(а)



(б)

Рисунок 25. Кривые обучения линейной регрессии в случае изменения погрешности смещения магнита вдоль оси y для десяти магнитов. (а) – в приближении, где количество обучающих примеров меньше 100, (б) – общие графики кривых обучения

Как видно по кривой для обучающего набора, в нем присутствуют больше 20 образцов, к которым модель в полной мере может быть подогнана (рис.25, а), что и объясняет начало кривой с ошибкой равной нулю. По мере добавления данных появляется нелинейность, ошибка ползет вверх, пока не стабилизируется (рис.25, б). По производительности модели на тестовом множестве можно сказать следующее: Обучение модели на маленьком количестве данных является не очень эффективным. Ошибка постоянно увеличивается. Как только обучающий набор становится максимально обобщенным, то ошибка уменьшается, и кривая приближается к кривой обучающего набора.

Таблица 19. Оценки качества модели полиномиальной регрессии в случае изменения погрешности смещения магнита вдоль оси y для десяти магнитов

| | RMSE | | R^2 | | MAE | |
|----------|------------------|------------------|-------|------|------------------|------------------|
| | train | test | train | test | train | test |
| M2A.1_DY | $2,60 * 10^{-5}$ | $2,67 * 10^{-5}$ | 0,93 | 0,93 | $2,08 * 10^{-5}$ | $2,07 * 10^{-5}$ |
| M2B.1_DY | $5,17 * 10^{-5}$ | $5,17 * 10^{-5}$ | 0,73 | 0,73 | $4,13 * 10^{-5}$ | $4,15 * 10^{-5}$ |

| | | | | | | |
|----------|------------------|------------------|------|------|------------------|------------------|
| M3A.1_DY | $7,80 * 10^{-5}$ | $7,86 * 10^{-5}$ | 0,40 | 0,37 | $6,30 * 10^{-5}$ | $6,30 * 10^{-5}$ |
| M3B.1_DY | $5,43 * 10^{-5}$ | $5,54 * 10^{-5}$ | 0,70 | 0,67 | $4,33 * 10^{-5}$ | $4,42 * 10^{-5}$ |
| M4A.1_DY | $8,23 * 10^{-5}$ | $8,55 * 10^{-5}$ | 0,32 | 0,32 | $6,55 * 10^{-5}$ | $6,83 * 10^{-5}$ |
| M4B.1_DY | $5,39 * 10^{-4}$ | $5,53 * 10^{-5}$ | 0,71 | 0,69 | $4,31 * 10^{-5}$ | $4,40 * 10^{-5}$ |
| M5A.1_DY | $8,23 * 10^{-5}$ | $8,25 * 10^{-5}$ | 0,32 | 0,31 | $6,56 * 10^{-5}$ | $6,56 * 10^{-5}$ |
| M5B.1_DY | $5,27 * 10^{-5}$ | $5,24 * 10^{-5}$ | 0,73 | 0,70 | $4,20 * 10^{-5}$ | $4,28 * 10^{-5}$ |
| M6A.1_DY | $8,32 * 10^{-5}$ | $8,32 * 10^{-5}$ | 0,32 | 0,32 | $6,64 * 10^{-5}$ | $6,67 * 10^{-5}$ |
| M6B.1_DY | $4,15 * 10^{-5}$ | $4,14 * 10^{-5}$ | 0,82 | 0,83 | $3,31 * 10^{-5}$ | $3,38 * 10^{-5}$ |

Исходя из результатов в таблицах 18,19 видно, что оценки качества для линейной и полиномиальной регрессий практически не отличаются друг от друга и примерно равны среднему значению смещения магнита вдоль оси u , соответственно, может существовать модель, более четко описывающая данные.

Затем была построена модель многослойного перцептрона. В обучении было проведено 200 эпох, на каждой из которых модель видит по 16 обучающих примера прежде, чем обновить веса. Число нейронов в одном слое равно 100, всего 3 скрытых слоя. Функция активации – softsign (формула 14). Функция потерь MSE и MAE. Ошибки обучения для каждого магнита можно увидеть в таблице 20.

Таблица 20. Оценки качества модели многослойного перцептрона с метриками MSE и MAE в случае изменения погрешности смещения магнита вдоль оси u для десяти магнитов

| | <i>MSE</i> | | <i>MAE</i> | |
|----------|------------------|------------------|------------------|------------------|
| | <i>train</i> | <i>test</i> | <i>train</i> | <i>test</i> |
| M2A.1_DY | $9,53 * 10^{-9}$ | $9,58 * 10^{-9}$ | $2,93 * 10^{-5}$ | $2,90 * 10^{-5}$ |
| M2B.1_DY | $5,06 * 10^{-9}$ | $5,14 * 10^{-9}$ | $2,03 * 10^{-5}$ | $2,03 * 10^{-5}$ |
| M3A.1_DY | $1,68 * 10^{-9}$ | $1,63 * 10^{-9}$ | $1,05 * 10^{-5}$ | $1,10 * 10^{-5}$ |
| M3B.1_DY | $1,09 * 10^{-9}$ | $1,14 * 10^{-9}$ | $8,35 * 10^{-6}$ | $8,31 * 10^{-6}$ |
| M4A.1_DY | $1,69 * 10^{-9}$ | $1,79 * 10^{-9}$ | $1,06 * 10^{-5}$ | $1,08 * 10^{-5}$ |
| M4B.1_DY | $1,01 * 10^{-9}$ | $1,03 * 10^{-9}$ | $7,96 * 10^{-6}$ | $8,03 * 10^{-6}$ |
| M5A.1_DY | $2,24 * 10^{-9}$ | $2,27 * 10^{-9}$ | $1,24 * 10^{-5}$ | $1,22 * 10^{-5}$ |

| | | | | |
|----------|-------------------|------------------|------------------|------------------|
| M5B.1_DY | $9,31 * 10^{-10}$ | $1,04 * 10^{-9}$ | $8,06 * 10^{-6}$ | $7,97 * 10^{-6}$ |
| M6A.1_DY | $3,54 * 10^{-9}$ | $3,81 * 10^{-9}$ | $1,64 * 10^{-5}$ | $1,72 * 10^{-5}$ |
| M6B.1_DY | $8,89 * 10^{-9}$ | $8,93 * 10^{-9}$ | $2,79 * 10^{-5}$ | $2,79 * 10^{-5}$ |

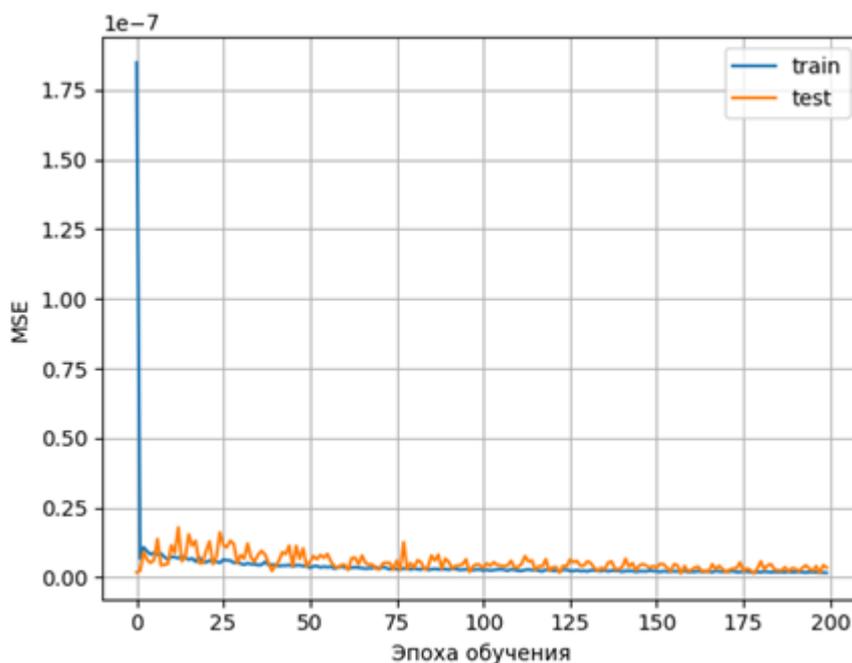


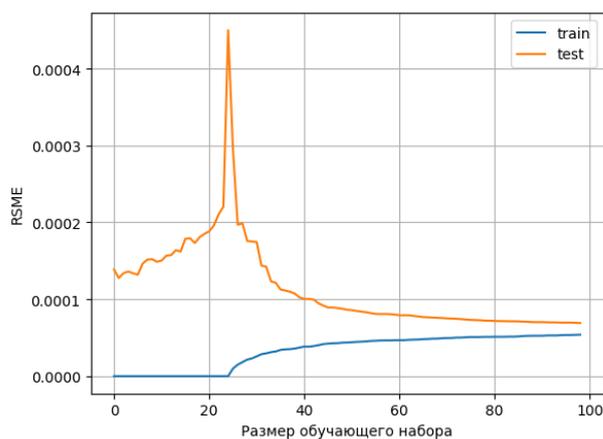
Рисунок 26. График зависимости среднеквадратичной ошибки от эпохи обучения модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси у для десяти магнитов

Как видно на рис.26, на первых эпохах обучения ошибка на обучающем множестве значительно превосходит ошибку на тестовом множестве. Для тестового множества видим колеблющиеся значения метрики чуть превосходящие значения на обучающем наборе.

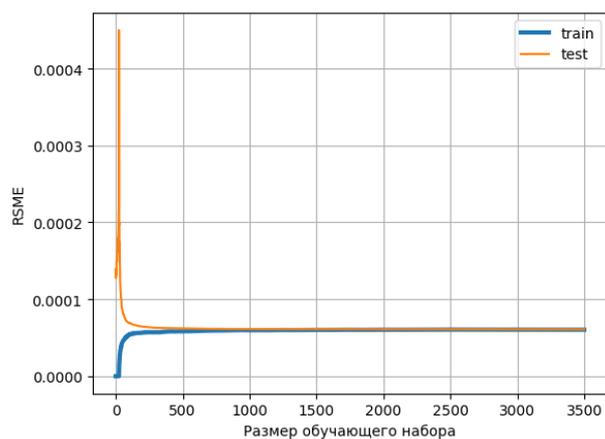
Для изменения погрешности смещения вдоль оси у десяти магнитов можно сказать, что модели линейной регрессии и многослойного перцептрона имеют приблизительно одинаковую оценку качества, которая по порядку величины равна самим значениям погрешности.

5.3.3. Сорок магнитов

По аналогии с предыдущими подпунктами для 40 магнитов были построены 2 вида моделей: линейная регрессия, модель многослойного перцептрона. Данные по линейной регрессии можно увидеть ниже.



(а)



(б)

Рисунок 27. Кривые обучения линейной регрессии в случае изменения погрешности смещения магнита вдоль оси у для сорока магнитов (а) – в приближении, где количество обучающих примеров меньше 100, (б) – общие графики кривых обучения

Таблица 21. Оценки качества модели линейной регрессии случае изменения погрешности смещения магнита вдоль оси у для сорока магнитов

| | RMSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DY | $2,59 * 10^{-5}$ | $2,57 * 10^{-5}$ | $2,07 * 10^{-5}$ | $2,07 * 10^{-5}$ |
| M2B.1_DY | $5,17 * 10^{-5}$ | $5,14 * 10^{-5}$ | $4,14 * 10^{-5}$ | $4,13 * 10^{-5}$ |
| M3A.1_DY | $7,86 * 10^{-5}$ | $7,82 * 10^{-5}$ | $6,29 * 10^{-5}$ | $6,29 * 10^{-5}$ |
| M3B.1_DY | $5,43 * 10^{-5}$ | $5,55 * 10^{-5}$ | $4,34 * 10^{-5}$ | $4,42 * 10^{-5}$ |
| M4A.1_DY | $8,21 * 10^{-5}$ | $8,57 * 10^{-5}$ | $6,54 * 10^{-5}$ | $6,84 * 10^{-5}$ |
| M4B.1_DY | $5,39 * 10^{-5}$ | $5,54 * 10^{-5}$ | $4,32 * 10^{-5}$ | $4,44 * 10^{-5}$ |
| M5A.1_DY | $8,22 * 10^{-5}$ | $8,24 * 10^{-5}$ | $6,56 * 10^{-5}$ | $6,59 * 10^{-5}$ |
| M5B.1_DY | $5,26 * 10^{-5}$ | $5,27 * 10^{-5}$ | $4,21 * 10^{-5}$ | $4,20 * 10^{-5}$ |
| M6A.1_DY | $8,32 * 10^{-5}$ | $8,33 * 10^{-5}$ | $6,64 * 10^{-5}$ | $6,69 * 10^{-5}$ |
| M6B.1_DY | $4,15 * 10^{-5}$ | $4,16 * 10^{-5}$ | $3,31 * 10^{-5}$ | $3,39 * 10^{-5}$ |
| M2A.2_DY | $2,52 * 10^{-5}$ | $2,58 * 10^{-5}$ | $2,03 * 10^{-5}$ | $2,15 * 10^{-5}$ |
| M2B.2_DY | $5,03 * 10^{-5}$ | $4,96 * 10^{-5}$ | $4,45 * 10^{-5}$ | $4,90 * 10^{-5}$ |
| M3A.2_DY | $7,69 * 10^{-5}$ | $7,66 * 10^{-5}$ | $5,97 * 10^{-5}$ | $6,17 * 10^{-5}$ |
| M3B.2_DY | $5,42 * 10^{-5}$ | $5,46 * 10^{-5}$ | $4,37 * 10^{-5}$ | $4,44 * 10^{-5}$ |
| M4A.2_DY | $8,23 * 10^{-5}$ | $5,15 * 10^{-5}$ | $6,57 * 10^{-5}$ | $6,59 * 10^{-5}$ |
| M4B.2_DY | $5,50 * 10^{-5}$ | $5,51 * 10^{-5}$ | $4,41 * 10^{-5}$ | $4,42 * 10^{-5}$ |
| M5A.2_DY | $8,38 * 10^{-5}$ | $8,37 * 10^{-5}$ | $6,74 * 10^{-5}$ | $6,78 * 10^{-5}$ |

| | | | | |
|----------|------------------|------------------|------------------|------------------|
| M5B.2_DY | $5,34 * 10^{-5}$ | $5,53 * 10^{-5}$ | $4,29 * 10^{-5}$ | $4,34 * 10^{-5}$ |
| M6A.2_DY | $8,13 * 10^{-5}$ | $8,54 * 10^{-5}$ | $6,52 * 10^{-5}$ | $6,75 * 10^{-5}$ |
| M6B.2_DY | $4,06 * 10^{-5}$ | $4,26 * 10^{-5}$ | $3,25 * 10^{-5}$ | $3,37 * 10^{-5}$ |
| M2A.3_DY | $2,52 * 10^{-5}$ | $2,53 * 10^{-5}$ | $2,01 * 10^{-5}$ | $2,04 * 10^{-5}$ |
| M2B.3_DY | $5,04 * 10^{-5}$ | $5,04 * 10^{-5}$ | $4,01 * 10^{-5}$ | $4,07 * 10^{-5}$ |
| M3A.3_DY | $7,27 * 10^{-5}$ | $7,68 * 10^{-5}$ | $6,17 * 10^{-5}$ | $6,21 * 10^{-5}$ |
| M3B.3_DY | $5,45 * 10^{-5}$ | $5,49 * 10^{-5}$ | $4,39 * 10^{-5}$ | $4,39 * 10^{-5}$ |
| M4A.3_DY | $8,12 * 10^{-5}$ | $8,26 * 10^{-5}$ | $6,48 * 10^{-5}$ | $6,58 * 10^{-5}$ |
| M4B.3_DY | $5,36 * 10^{-5}$ | $5,50 * 10^{-5}$ | $4,32 * 10^{-5}$ | $4,43 * 10^{-5}$ |
| M5A.3_DY | $8,31 * 10^{-5}$ | $8,30 * 10^{-5}$ | $6,64 * 10^{-5}$ | $6,61 * 10^{-5}$ |
| M5B.3_DY | $5,31 * 10^{-5}$ | $5,34 * 10^{-5}$ | $4,21 * 10^{-5}$ | $4,24 * 10^{-5}$ |
| M6A.3_DY | $8,20 * 10^{-5}$ | $8,22 * 10^{-5}$ | $6,50 * 10^{-5}$ | $6,51 * 10^{-5}$ |
| M6B.3_DY | $4,09 * 10^{-5}$ | $4,10 * 10^{-5}$ | $3,24 * 10^{-5}$ | $3,25 * 10^{-5}$ |
| M2A.4_DY | $2,52 * 10^{-5}$ | $2,56 * 10^{-5}$ | $2,00 * 10^{-5}$ | $2,04 * 10^{-5}$ |
| M2B.4_DY | $5,03 * 10^{-5}$ | $5,11 * 10^{-5}$ | $4,00 * 10^{-5}$ | $4,08 * 10^{-5}$ |
| M3A.4_DY | $7,63 * 10^{-5}$ | $7,72 * 10^{-5}$ | $6,05 * 10^{-5}$ | $6,18 * 10^{-5}$ |
| M3B.4_DY | $5,37 * 10^{-5}$ | $5,41 * 10^{-5}$ | $4,28 * 10^{-5}$ | $4,36 * 10^{-5}$ |
| M4A.4_DY | $8,26 * 10^{-5}$ | $8,34 * 10^{-5}$ | $6,75 * 10^{-5}$ | $6,64 * 10^{-5}$ |
| M4B.4_DY | $5,52 * 10^{-5}$ | $5,52 * 10^{-5}$ | $4,41 * 10^{-5}$ | $4,46 * 10^{-5}$ |
| M5A.4_DY | $8,40 * 10^{-5}$ | $8,51 * 10^{-5}$ | $6,72 * 10^{-5}$ | $6,80 * 10^{-5}$ |
| M5B.4_DY | $5,37 * 10^{-5}$ | $5,45 * 10^{-5}$ | $4,26 * 10^{-5}$ | $4,25 * 10^{-5}$ |
| M6A.4_DY | $8,16 * 10^{-5}$ | $8,26 * 10^{-5}$ | $6,52 * 10^{-5}$ | $6,55 * 10^{-5}$ |
| M6B.4_DY | $4,07 * 10^{-5}$ | $4,12 * 10^{-5}$ | $3,25 * 10^{-5}$ | $3,31 * 10^{-5}$ |

По результатам описанным в таблице 21 можно сказать, что линейная регрессия неплохо описывает данные из датасета.

Построим модель многослойного перцептрона со следующими параметрами: 4 слоя + 1 выходной, количество нейронов в каждом слое 160, функция активации softsign (формула 14). Оптимизатор модели – adam. В качестве метрик были взяты MSE и MAE. Графики изменения ошибок от эпохи обучения представлен на рис.28, рис.29.

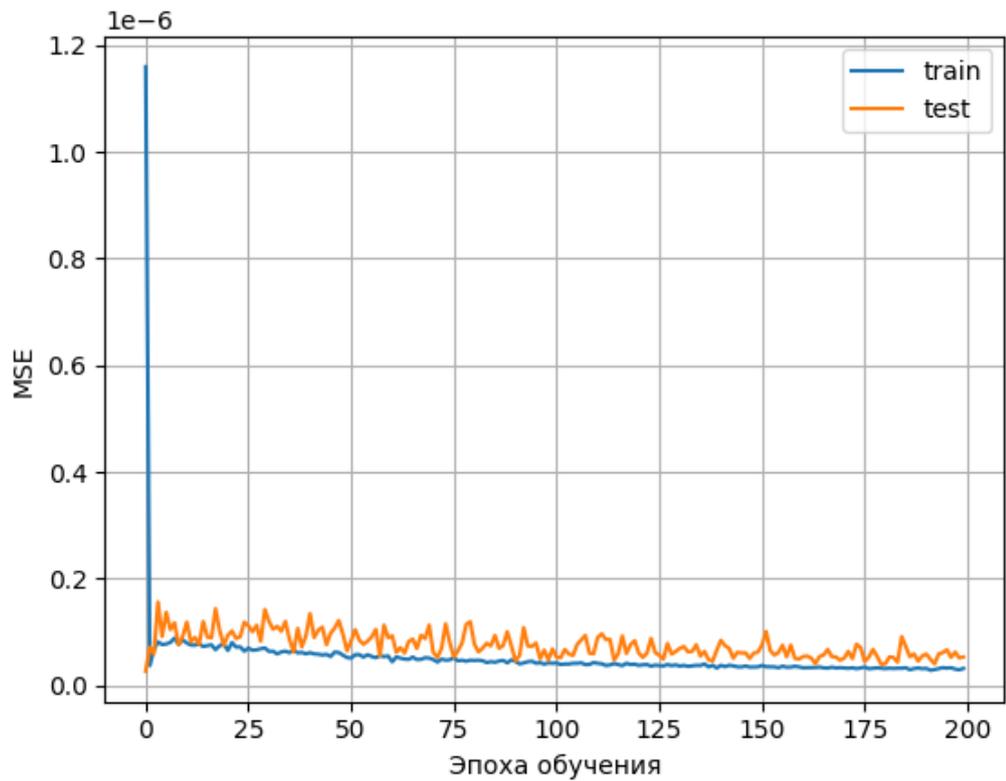


Рисунок 28. График зависимости метрики MSE от эпохи обучения для модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси u для сорока магнитов.

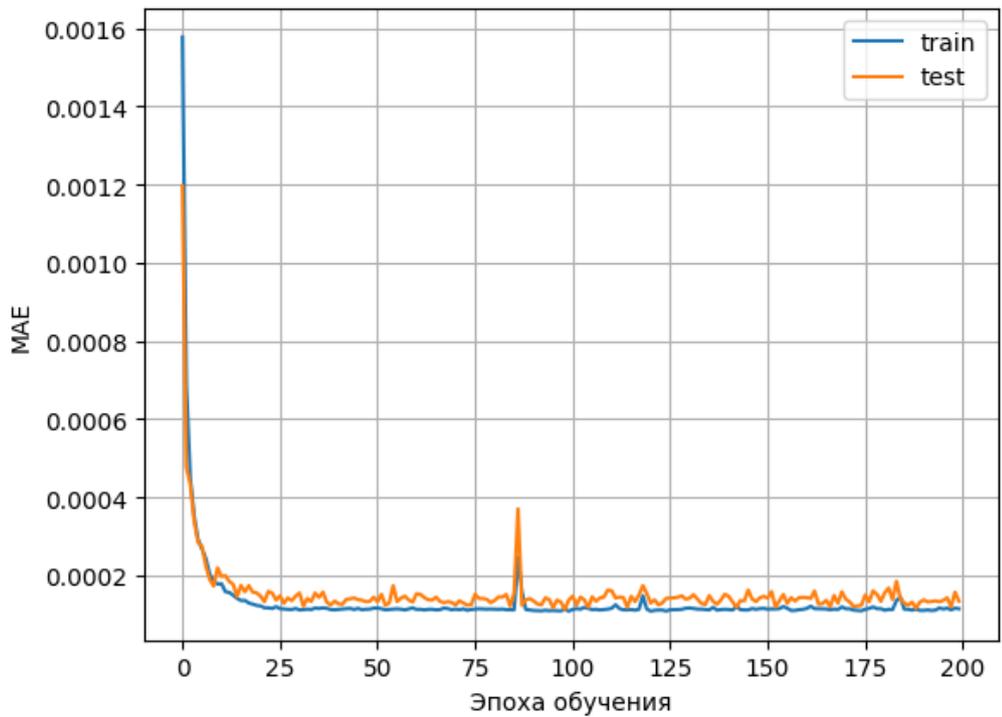


Рисунок 29. График зависимости метрики MAE от эпохи обучения для модели многослойного перцептрона в случае изменения погрешности смещения вдоль оси u для сорока магнитов.

Таблица 22. Оценки качества модели многослойного перцептрона с метриками MSE и MAE в случае изменения погрешности смещения магнита вдоль оси у для сорока магнитов

| | RMSE | | MAE | |
|----------|------------------|------------------|------------------|------------------|
| | train | test | train | test |
| M2A.1_DY | $1,07 * 10^{-4}$ | $1,15 * 10^{-4}$ | $9,16 * 10^{-5}$ | $9,21 * 10^{-5}$ |
| M2B.1_DY | $2,78 * 10^{-4}$ | $2,78 * 10^{-4}$ | $9,55 * 10^{-5}$ | $9,50 * 10^{-5}$ |
| M3A.1_DY | $1,60 * 10^{-4}$ | $1,62 * 10^{-4}$ | $8,61 * 10^{-5}$ | $8,71 * 10^{-5}$ |
| M3B.1_DY | $1,07 * 10^{-4}$ | $1,04 * 10^{-4}$ | $1,31 * 10^{-4}$ | $1,31 * 10^{-4}$ |
| M4A.1_DY | $3,76 * 10^{-4}$ | $3,78 * 10^{-4}$ | $3,75 * 10^{-4}$ | $3,77 * 10^{-4}$ |
| M4B.1_DY | $2,37 * 10^{-4}$ | $2,43 * 10^{-4}$ | $9,68 * 10^{-5}$ | $9,68 * 10^{-5}$ |
| M5A.1_DY | $3,01 * 10^{-4}$ | $3,02 * 10^{-4}$ | $1,87 * 10^{-4}$ | $1,86 * 10^{-4}$ |
| M5B.1_DY | $1,59 * 10^{-4}$ | $1,57 * 10^{-4}$ | $9,77 * 10^{-5}$ | $1,02 * 10^{-4}$ |
| M6A.1_DY | $1,57 * 10^{-4}$ | $1,51 * 10^{-4}$ | $7,98 * 10^{-5}$ | $8,04 * 10^{-5}$ |
| M6B.1_DY | $1,94 * 10^{-4}$ | $1,95 * 10^{-4}$ | $9,00 * 10^{-5}$ | $9,11 * 10^{-5}$ |
| M2A.2_DY | $1,06 * 10^{-4}$ | $1,12 * 10^{-4}$ | $9,56 * 10^{-5}$ | $9,52 * 10^{-5}$ |
| M2B.2_DY | $1,56 * 10^{-4}$ | $1,63 * 10^{-4}$ | $7,89 * 10^{-5}$ | $8,01 * 10^{-5}$ |
| M3A.2_DY | $1,99 * 10^{-4}$ | $2,03 * 10^{-4}$ | $9,97 * 10^{-5}$ | $1,03 * 10^{-4}$ |
| M3B.2_DY | $1,19 * 10^{-4}$ | $1,19 * 10^{-4}$ | $8,13 * 10^{-5}$ | $8,20 * 10^{-5}$ |
| M4A.2_DY | $3,82 * 10^{-4}$ | $3,88 * 10^{-4}$ | $1,06 * 10^{-4}$ | $1,06 * 10^{-4}$ |
| M4B.2_DY | $2,96 * 10^{-4}$ | $2,99 * 10^{-4}$ | $1,26 * 10^{-4}$ | $1,32 * 10^{-4}$ |
| M5A.2_DY | $1,00 * 10^{-4}$ | $1,01 * 10^{-5}$ | $8,09 * 10^{-5}$ | $8,16 * 10^{-5}$ |
| M5B.2_DY | $9,83 * 10^{-4}$ | $1,00 * 10^{-5}$ | $1,14 * 10^{-4}$ | $1,11 * 10^{-4}$ |
| M6A.2_DY | $1,43 * 10^{-4}$ | $1,46 * 10^{-4}$ | $1,90 * 10^{-4}$ | $1,90 * 10^{-4}$ |
| M6B.2_DY | $3,06 * 10^{-4}$ | $3,13 * 10^{-4}$ | $3,26 * 10^{-5}$ | $3,34 * 10^{-4}$ |
| M2A.3_DY | $2,10 * 10^{-4}$ | $2,18 * 10^{-4}$ | $8,03 * 10^{-5}$ | $7,99 * 10^{-5}$ |
| M2B.3_DY | $1,05 * 10^{-4}$ | $1,09 * 10^{-4}$ | $2,10 * 10^{-4}$ | $2,18 * 10^{-4}$ |
| M3A.3_DY | $2,24 * 10^{-4}$ | $2,24 * 10^{-4}$ | $9,62 * 10^{-5}$ | $9,66 * 10^{-5}$ |
| M3B.3_DY | $1,81 * 10^{-4}$ | $1,78 * 10^{-4}$ | $8,15 * 10^{-5}$ | $8,21 * 10^{-5}$ |
| M4A.3_DY | $1,78 * 10^{-4}$ | $1,77 * 10^{-4}$ | $1,51 * 10^{-4}$ | $1,50 * 10^{-4}$ |
| M4B.3_DY | $1,58 * 10^{-4}$ | $1,60 * 10^{-4}$ | $1,26 * 10^{-4}$ | $1,26 * 10^{-4}$ |
| M5A.3_DY | $2,04 * 10^{-4}$ | $2,00 * 10^{-4}$ | $9,64 * 10^{-5}$ | $9,65 * 10^{-5}$ |
| M5B.3_DY | $1,19 * 10^{-4}$ | $1,24 * 10^{-4}$ | $8,10 * 10^{-5}$ | $8,41 * 10^{-5}$ |

| | | | | |
|----------|------------------|------------------|------------------|------------------|
| M6A.3_DY | $1,16 * 10^{-4}$ | $1,21 * 10^{-4}$ | $2,72 * 10^{-4}$ | $2,72 * 10^{-4}$ |
| M6B.3_DY | $1,59 * 10^{-4}$ | $1,59 * 10^{-4}$ | $8,35 * 10^{-5}$ | $8,46 * 10^{-5}$ |
| M2A.4_DY | $1,81 * 10^{-4}$ | $1,81 * 10^{-4}$ | $1,36 * 10^{-4}$ | $1,37 * 10^{-4}$ |
| M2B.4_DY | $2,92 * 10^{-4}$ | $2,90 * 10^{-4}$ | $2,62 * 10^{-4}$ | $2,60 * 10^{-4}$ |
| M3A.4_DY | $1,27 * 10^{-4}$ | $1,28 * 10^{-4}$ | $7,98 * 10^{-5}$ | $8,25 * 10^{-5}$ |
| M3B.4_DY | $1,33 * 10^{-4}$ | $1,36 * 10^{-4}$ | $8,38 * 10^{-5}$ | $8,33 * 10^{-5}$ |
| M4A.4_DY | $1,65 * 10^{-4}$ | $1,80 * 10^{-4}$ | $1,30 * 10^{-4}$ | $1,36 * 10^{-4}$ |
| M4B.4_DY | $2,48 * 10^{-4}$ | $2,45 * 10^{-4}$ | $9,87 * 10^{-5}$ | $9,92 * 10^{-5}$ |
| M5A.4_DY | $2,54 * 10^{-4}$ | $2,53 * 10^{-4}$ | $8,58 * 10^{-5}$ | $8,86 * 10^{-5}$ |
| M5B.4_DY | $1,07 * 10^{-4}$ | $1,09 * 10^{-4}$ | $1,14 * 10^{-4}$ | $1,17 * 10^{-4}$ |
| M6A.4_DY | $1,05 * 10^{-4}$ | $1,07 * 10^{-4}$ | $1,08 * 10^{-4}$ | $1,05 * 10^{-4}$ |
| M6B.4_DY | $3,05 * 10^{-4}$ | $3,05 * 10^{-4}$ | $8,03 * 10^{-5}$ | $7,94 * 10^{-5}$ |

Как видно из значений в таблице 22 ошибка на тестовом и обучающем множествах при построении модели многослойного перцептрона для 40 магнитов имеет примерно такой же порядок, что и сами значения предсказываемого параметра, а соответственно полученное значение может не попадать в допустимый предел погрешности 20%. Исходя из показателей ошибки MAE, модель линейной регрессии чуть лучше описывает наши данные, чем модель многослойного перцептрона.

Заключение

В данной работе были изучены такие погрешности параметров магнитов как смещение магнита вдоль осей x , y (DX , DY), а также погрешность угла поворота магнита вокруг оси распространения пучка ($DPSI$).

При помощи языка программирования Python был написан код для формирования датасетов, состоящих из погрешностей для n магнитов и компонент траектории пучка заряженных частиц.

Для каждого параметра магнита были построены модели машинного обучения, такие как линейная регрессия, полиномиальная регрессия, многослойный перцептрон. Были определены оценки качества.

При изменении погрешностей смещения на одном или двух магнитах линейные и полиномиальные регрессии показали высокую оценку качества этих моделей, что позволило ограничиться только ими. В случае, варьирования погрешностей на десяти магнитах все модели, включая многослойный перцептрон, выдали приблизительно одинаковый показатель качества. Для сорока магнитов наилучшей моделью, описывающей имеющиеся данные, оказалась линейная регрессия.

Список используемых сокращений

МэВ – мегаэлектронвольт

Циклотрон – циклический ускоритель

μ -мезоны – мюоны

π -мезоны – пионы

ОИЯИ – Объединенный институт ядерных исследований в Дубне

NICA – Nuclotron-based Ion Collider facility- установка на встречных пучках, на основе ускорителя нуклотрон

ЦЕРН – европейская организация по ядерным исследованиям, крупнейшая лаборатория физики высоких энергий

ВЧ резонансный ускоритель – Высокочастотный резонансный ускоритель

ESFRI – Европейская организация по стратегическому развитию исследовательской инфраструктуры

АЭС – атомная электростанция

МО – машинное обучение

Train – обучающее множество

Test – тестовое множество

RMSE – корень из среднеквадратичной ошибки

MAE – среднее абсолютное значение ошибок

Список литературы

1. Комар Е.Г. Основы ускорительной техники. Москва: Атом-издат, 1975. 368 с.
2. NICA. Nuclotron-based Ion Collider fAcility [Электронный ресурс]. URL: <https://nica.jinr.ru/complex.php>.
3. Вьюгин В.В. Математические основы машинного обучения и прогнозирования. МОсква, 2018. 484 с.
4. Кравец С.Л. Ускорители заряженных частиц [Электронный ресурс] // Научно-образовательный портал «Большая российская энциклопедия». URL: <https://bigenc.ru/c/uskoriteli-zariazhennykh-chastits-1d2f47>.
5. Лебедев А.Н., Шальнов А.. Основы физики и техники ускорителей. Т.1. Ускорители заряженных частиц. Москва: Энергоиздат, 1981. 192 с.
6. Черняев А.П., Белихин М.А., Желтоножская М.В. Введение в физику ускорителей заряженных частиц. 2019.
7. Бурштейн Э.Л. Ускорители заряженных частиц [Электронный ресурс]. URL: <https://www.booksite.ru/fulltext/1/001/008/114/616.htm>.
8. Векслер В.И. Ускорители атомных частиц [Электронный ресурс] // Научно-образовательный портал «Большая российская энциклопедия». URL: http://elib.biblioatom.ru/text/veksler_uskoriteli-atomnyh-chastits_1956/go,16/.
9. Сидорин А.О. Ускорители заряженных частиц [Электронный ресурс] // Большая Российская энциклопедия. URL: <https://old.bigenc.ru/physics/text/4702177>.
10. Будкер Г.И. Ускорители со встречными пучками // Успехи физических наук, том 89 выпуск 4. 1966. С. 16.
11. CERN. The Higgs boson [Электронный ресурс] // CERN. 2012. URL:

- <https://home.cern/science/physics/higgs-boson>.
12. Open Education at JINR [Электронный ресурс]. URL: <https://edu.jinr.ru/>.
 13. Ускорители заряженных частиц. Классификация ускорителей [Электронный ресурс] // Большая советская энциклопедия. URL: <http://www.help-rus-student.ru/text/82/641.htm>.
 14. Исследования О. институт ядерных. Краткий отчето созданиии и эксплуатации комплекса сверхпроводящих колец на встречных пучках тяжелых ионов NICA (Комплекса NICA). 2020. 1–196 с.
 15. Мешков И.Н., Требников Г.В., Сидорин А.О. В составлении Исходных данных принимали участие : Дубна, 2013. С. 90.
 16. Ziemann V. Hands on accelerator physics using Matlab. CRC Press is an imprint of Taylor & Francis Group, an Informa business, 2019.
 17. Iselin F.C. The MadX Program (Methodical Accelerator Design): Reference Manual. 1985. Т. 13.
 18. Галушкин А.И. Нейронные сети [Электронный ресурс]. URL: https://old.bigenc.ru/technology_and_technique/text/4114009.
 19. Evangreen [Электронный ресурс]. URL: <https://evergreens.com.ua/ru/articles/machine-learning-overview.html>.
 20. Салтунов Г. Введение в машинное обучение.
 21. Geron A. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow концепции, инструменты и техники для создания интеллектуальн.. / под ред. Артеменко Т.Н., В.Г. П. Beijing, Boston, Farnham, Sebastopol, Tokyo, 2018. 688 с.
 22. The python API for MAD-X [Электронный ресурс]. URL: <https://pymad.github.io/cpymad/cpymad/madx.html>.
 23. Джулли А., Пал С. Библиотека Keras-инструмент глубокого обучения //

- М.: ДМК Пресс. 2018.
24. Marius-Constantin P. и др. Multilayer perceptron and neural networks // WSEAS Trans. Circuits Syst. 2009. Т. 8, № 7. С. 579–588.
 25. Neural network models (supervised) [Электронный ресурс] // siklit-learn. URL: https://scikit-learn.org/stable/modules/neural_networks_supervised.html.
 26. Python [Электронный ресурс]. URL: <https://wiki.python.org/moin/BeginnersGuide/Overview>.
 27. Ruder S. An overview of gradient descent optimization algorithms. 2016. С. 1–14.