

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА»

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА ТЕОРИИ ВЕРОЯТНОСТЕЙ

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
специалиста

**ВОССТАНОВЛЕНИЕ НЕЙТРАЛЬНЫХ ЛИНИЙ МАГНИТНОГО ПОЛЯ
СОЛНЦА МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ**

Выполнил студент 609 группы
Киселюс Вацловас

подпись студента

Научный руководитель:
к.ф.-м.н., доцент
Илларионов Егор Александрович

подпись научного руководителя

Москва
2023 год

Содержание

1	Введение	2
1.1	Солнечная активность	2
1.2	Синоптические карты	2
1.3	Постановка задачи	2
1.4	Цель работы	3
2	Предлагаемый подход	4
2.1	Идея	4
2.2	Описание подхода	5
2.2.1	Полносвязная нейронная сеть	5
2.2.2	Функция потерь	6
3	Датасет	6
3.1	Предпосылки к созданию датасета	6
3.2	Вероятностный подход	6
3.2.1	Введение	6
3.2.2	Регрессия на основе гауссовских процессов	7
3.2.3	Применение в наших целях	8
3.3	Примеры сгенерированных изображений	9
4	Эксперимент на датасете	9
4.1	Структура эксперимента	9
4.2	Метрика качества	10
4.3	Результаты	11
4.3.1	Полная информация о нейтральной линии	11
4.3.2	50% информации о нейтральной линии	12
4.3.3	35% информации о нейтральной линии	13
4.3.4	20% информации о нейтральной линии	14
4.3.5	10% информации о нейтральной линии	15
4.4	Выводы из эксперимента	16
5	Эксперимент на реальных картах	17
6	Направления дальнейших исследований	18
6.1	Последовательные карты	18
6.2	Многомасштабный подход	18
7	Выводы	19
8	Заключение	19
	Список литературы	20

1 Введение

1.1 Солнечная активность

На Солнце постоянно происходят изменения, которые мы наблюдаем в виде солнечных пятен, корональных дыр, протуберанцев (см. рис. 1) и ряда других активных областей. Согласно современным представлениям, причиной образования активных областей являются магнитные поля Солнца. Предполагается, что протуберанцы (волокна) появляются вдоль нейтральных линий¹ магнитного поля Солнца.



Рис. 1: Протуберанец, или солнечное волокно.

Анализ наблюдательных данных по солнечной активности является основой для понимания процессов, протекающих на Солнце, и также для прогноза космической погоды [1, 2, 3].

1.2 Синоптические карты

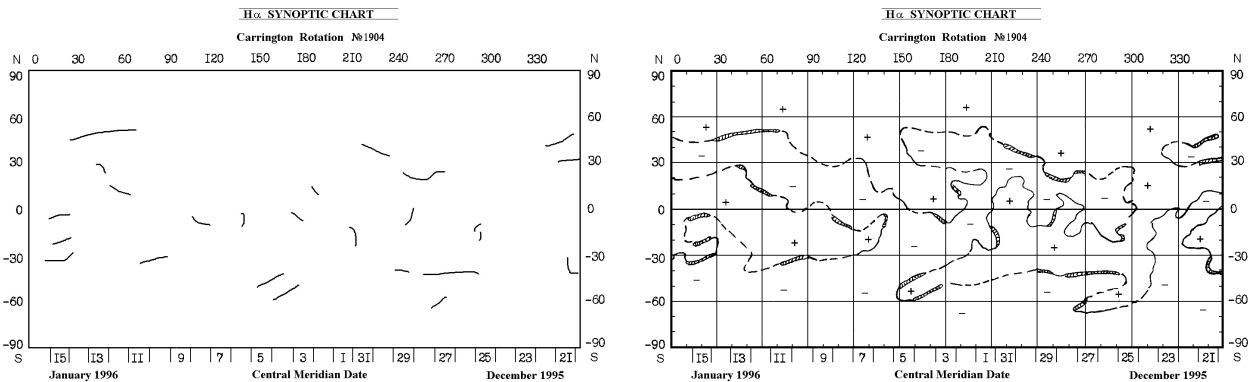


Рис. 2: Информация о волокнах, полученная за полный оборот Солнца вокруг своей оси и соответствующая синоптическая карта.

Одним из способов изучения солнечной активности является построение карт крупномасштабного магнитного поля, на которых, в частности, представляет интерес положение нейтральной линии магнитного поля (см. рис. 2). Существует несколько подходов [3, 4, 5] к восстановлению положения нейтральной линии и одним из них является восстановление линии по положениям волокон, задающим фрагменты этой линии [4]. Зачастую, эта процедура реализуется вручную, а результат оказывается неоднозначным, поскольку разные наблюдатели могут получать различные результаты для одних и тех же входных данных. Возникает вопрос, можно ли решить эту задачу методами машинного обучения?

1.3 Постановка задачи

Входными данными является информация о расположении волокон. Она представлена в виде матрицы $(a_{ij})_{m \times n}$, где m и n – размеры синоптической карты, i и j – соответствующие координаты. На выходе требуется получить матрицу $(b_{ij})_{m \times n}$, содержащую в себе

¹Т.е. на границе участков поверхности Солнца, знак магнитного поля которых отличается.

информацию о распределении знака магнитного поля и задающую тем самым положение нейтральной линии.

$$a_{ij} = \mathbb{1}\{\text{пиксель с координатами } (i, j) \text{ принадлежит волокну}\}$$

$$b_{ij} = 2 \cdot \mathbb{1}\{\text{знак поля положителен в пикселе с координатами } (i, j)\} - 1$$

Важно отметить, что левый и правый края каждого изображения отождествлены, так как мы имеем в виду проекцию поверхности Солнца на плоскость.

Вообще говоря, можно сформулировать задачу и вне контекста физики Солнца. Имея фрагменты некоторого числа замкнутых кривых (см. рис. 3а), необходимо соединить эти фрагменты в замкнутые кривые (см. рис. 3с). Для наглядности рассмотрим простой пример.

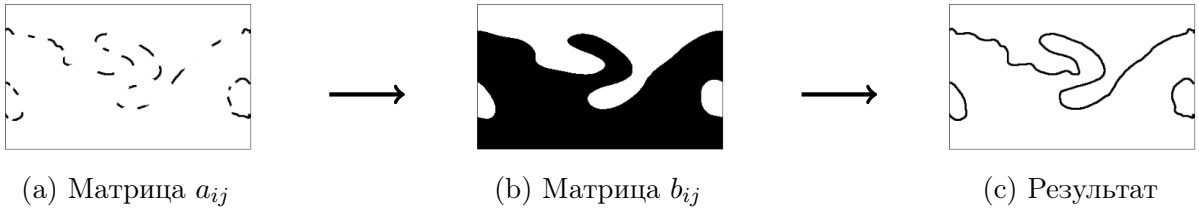


Рис. 3: Визуализация входных и выходных данных.

Как видно из формулировки, задача имеет множество решений. Однако, следующие факты помогут нам сузить область поиска возможных решений:

- волокна появляются вдоль нейтральных линий;
- вне нейтральной линии знак поля имеет знак $+$ или $-$;
- нам известен знак поля на северном и южном полюсах²;
- магнитное поле Солнца в целом нейтрально³.

1.4 Цель работы

Имеющиеся на сегодняшний день алгоритмы, основанные на эвристиках, работают неустойчиво. Целью работы является разработка подхода, основанного на машинном обучении. Кроме того, специфика задачи допускает, что решение может быть не единственно. Наш алгоритм должен предлагать множество решений, согласованных с наблюдательными данными.

²Т.к. в течение своего цикла активности (≈ 11 лет) знак поля на полюсах не меняется.

³«Плюсы» и «минусы» занимают примерно одинаковую площадь, т.е. значение $\sum_{(i,j)} b_{ij}$ близко к нулю.

2 Предлагаемый подход

2.1 Идея

Для того чтобы отождествить левый и правый края изображения (см. рис. 4а), предлагается поместить изображение на поверхность цилиндра \mathcal{C} так, чтобы нужные края совместились (см. рис. 4б).

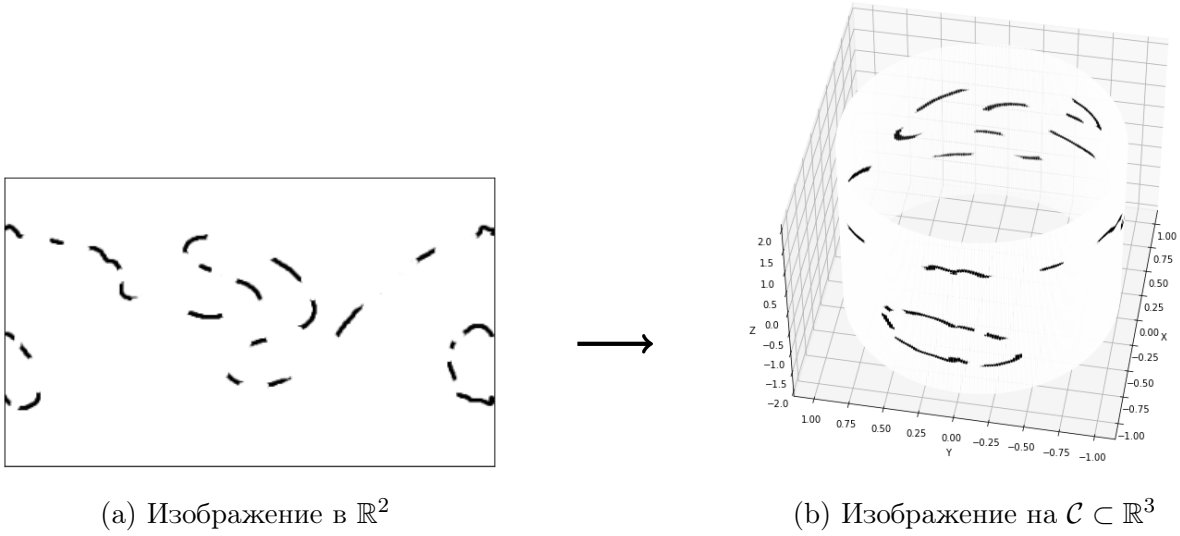


Рис. 4: Переход в трехмерное пространство.

Теперь можно попытаться построить функцию $f(x, y, z) : \mathcal{C} \subset \mathbb{R}^3 \rightarrow [-1, 1]$, для которой линия уровня ноль есть искомая нейтральная линия. На рис. 5 можно увидеть пример такой функции. Заметим, что зеленые *замкнутые кривые* в точности проходят вдоль волокон. По виду функции можно сказать, что вне нейтральной линии функция принимает значения ± 1 . Для построения таких функций необходимо сформулировать замечания из главы 1.3 на языке функций:

- $f|_{\mathcal{F}} = 0$, где \mathcal{F} – область, соответствующая наблюдаемым волокнам;
- $f|_{\mathcal{C} \setminus \mathcal{F}} = \pm 1$;
- известны значения f вблизи полюсов;
- $\int_{\mathcal{C}} f dV = 0$;
- при переходе через границу (волокно, линию уровня ноль) функция меняет знак.

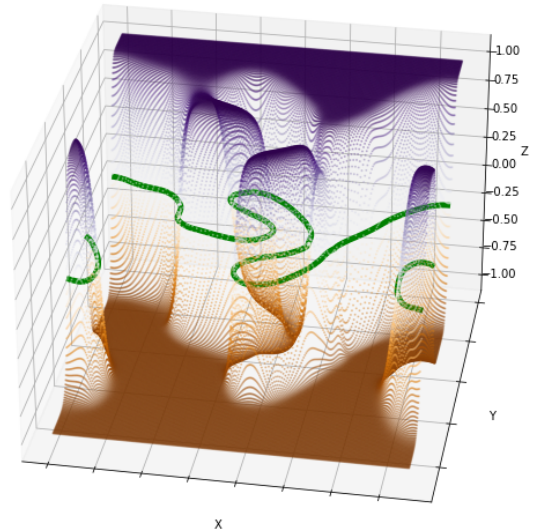


Рис. 5: Визуализация функции $f(x, y, z)|_{\mathcal{C}}$. Зеленым цветом отмечена линия уровня ноль.

Но как построить такую удобную функцию?

2.2 Описание подхода

2.2.1 Полносвязная нейронная сеть

Хорошо известно, что с помощью нейронных сетей можно конструировать сколь угодно сложные функции, они являются универсальными аппроксиматорами [6].

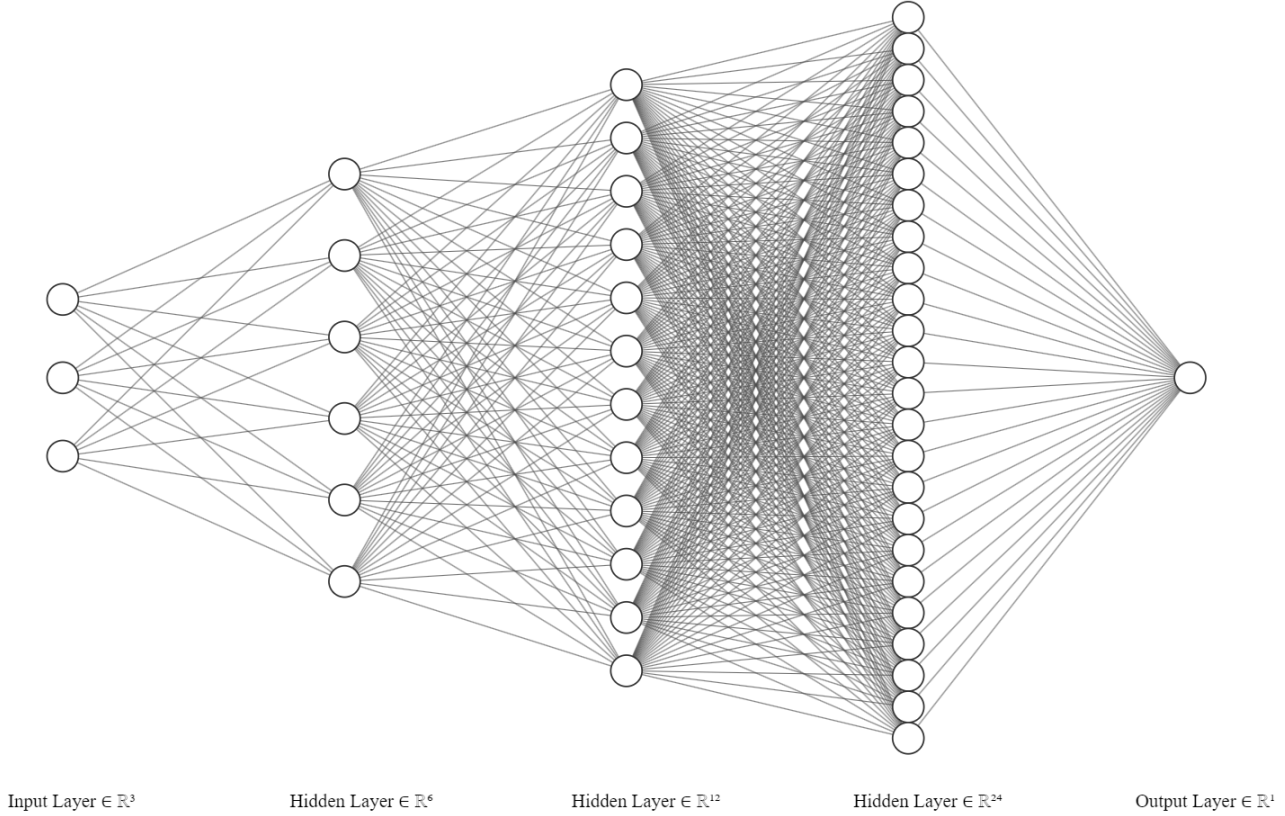


Рис. 6: Архитектура нейронной сети.

Например, мы можем использовать полносвязную нейронную сеть для решения сложных задач восстановления данных. В данном случае выбрана полносвязная нейронная сеть (см. рис. 6) с тремя входами (координаты точки на $\mathcal{C} \in \mathbb{R}^3$), тремя скрытыми слоями по 6, 12 и 24 нейрона в каждом слое, и выходной слой, состоящий из одного нейрона.

Нейронные сети используют функции активации для передачи сигнала от одного нейрона к другому. В данном случае для последнего слоя был выбран гиперболический тангенс, который имеет область значений $(-1, 1)$.

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Таким образом, мы можем интерпретировать выходное значение как знак магнитного поля в конкретном пикселе.

Для остальных слоев применяется функция ELU (exponential linear unit).

$$ELU(x) = \begin{cases} x & x \geq 0 \\ e^x - 1 & x < 0 \end{cases}$$

Перейдем к описанию loss-функции, используемой при обучении модели.

2.2.2 Функция потерь

Суть обучения нейронной сети заключается в том, что модель корректирует свои внутренние параметры (веса) так, чтобы минимизировался некоторый функционал, называемый *функцией потерь*, или *loss-функцией*. Поэтому одной из ключевых задач было создание функции потерь, необходимой для обучения модели, которая бы учитывала множество факторов, влияющих на форму нейтральной линии.

Исходя из рассуждений в главе 2.1, предлагается использовать следующую функцию потерь:

$$\mathcal{L}(f) = w_1 \left| \int_{\mathcal{C}} f dV \right| + w_2 \int_{\mathcal{F}} f^2 dV + \left(1 - w_3 \int_{\mathcal{C} \setminus \mathcal{F}} |f| dV \right) + \gamma$$

Здесь:

- $f = f(x, y, z)$ – функция, задаваемая нейронной сетью;
- w_i – нормировочные веса, подобранные так, чтобы каждое слагаемое в \mathcal{L} лежало в пределах от 0 до 1;
- \mathcal{C} – цилиндр, на который отображается входное изображение;
- \mathcal{F} – область, соответствующая наблюдаемым волокнам;
- член γ отвечает за ориентацию магнитного поля⁴.

3 Датасет

3.1 Предпосылки к созданию датасета

После реализации идеи в виде кода на языке *python* и проведения тестов на нарисованных от руки изображениях (например то, что использовалось ранее на рисунках) стало понятно, что подход рабочий. Однако оценка результатов проводилась «на глаз», что не могло являться аргументом в пользу состоятельности подхода. К тому же, необходимо иметь не только изображение с неполной границей, но и восстановленное. Сравнение предсказания модели с картами астрономов тоже не удовлетворяет нашим целям в виду ранее описанных причин.

Поэтому для тестирования модели была разработана процедура генерации синтетических изображений, которые похожи на возможное распределение магнитного поля и задают положение нейтральной линии. Ключевой идеей стало использование случайного трехмерного гауссовского поля.

3.2 Вероятностный подход

3.2.1 Введение

При обучении с учителем [7] часто используются параметрические модели $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$, чтобы смоделировать истинное распределение данных и найти оптимальные значения параметров $\boldsymbol{\theta}$ с помощью оценки максимального правдоподобия или оценка апостериорного

⁴Например, если заранее известно, что знак магнитного поля на северном полюсе отрицательный, а на южном – положительный, то мы требуем, чтобы функция принимала значения -1 и $+1$ на точках, принадлежащих верхнему и нижнему основаниям цилиндра \mathcal{C} соответственно. γ содержит в себе информацию о выполнении этого условия.

максимума [8]. Если необходимо, мы также можем вычислить апостериорное распределение $p(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})$ вместо точечной оценки $\hat{\boldsymbol{\theta}}$. С увеличением сложности и объема данных обычно требуются модели с большим числом параметров. Методы, использующие модели с фиксированным числом параметров, называются параметрическими методами.

В непараметрических методах количество параметров зависит от размера выборки. Например, в ядерной регрессии Надарая – Уотсона [9, 10] каждой наблюдаемой переменной y_i присваивается вес w_i и для прогнозирования значения в новой точке \mathbf{x} вычисляется средневзвешенное значение:

$$f(\mathbf{x}) = \sum_{i=1}^N w_i(\mathbf{x}) y_i$$

$$w_i(\mathbf{x}) = \frac{\kappa(\mathbf{x}, \mathbf{x}_i)}{\sum_{i'=1}^N \kappa(\mathbf{x}, \mathbf{x}_{i'})}$$

Наблюдения, которые находятся ближе к \mathbf{x} , имеют больший вес, чем наблюдения, которые находятся дальше. Веса вычисляются по основе \mathbf{x} и наблюдаемых \mathbf{x}_i с ядром κ . Частным случаем является алгоритм k ближайших соседей (KNN, [11]), где k ближайших наблюдений имеют вес $1/k$, а все остальные имеют вес 0. Непараметрические методы часто нуждаются в обработке всех обучающих данных для предсказания и поэтому медленнее по времени вывода, чем параметрические методы. С другой стороны, обучение обычно проходит быстрее, так как непараметрическим моделям необходимо запоминать только обучающие данные.

Еще одним примером непараметрических методов являются гауссовы процессы. Вместо того, чтобы выводить распределение по параметрам параметрической функции, гауссовы процессы могут использоваться для вывода распределения непосредственно по функциям. Гауссовский процесс определяет априорное распределение по функциям. После наблюдения некоторых значений функции его можно преобразовать в апостериорное распределение по функциям. Вывод значений функции в этом контексте называется регрессией на основе гауссовских процессов, или кригингом [12].

3.2.2 Регрессия на основе гауссовских процессов

Гауссовский процесс – это случайный процесс [13], в котором каждой точке $\mathbf{x} \in \mathbb{R}^d$ сопоставляется случайная переменная $f(\mathbf{x})$, а совместное распределение конечного числа этих переменных $p(f(\mathbf{x}_1), \dots, f(\mathbf{x}_N))$ является гауссовым.

$$p(\mathbf{f}|\mathbf{x}) = \mathcal{N}(\mathbf{f}|\boldsymbol{\mu}, \mathbf{K}) \quad (1)$$

В уравнении (1) $\mathbf{f} = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_N))$, $\boldsymbol{\mu} = (m(\mathbf{x}_1), \dots, m(\mathbf{x}_N))$, а $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$. Функция m является функцией среднего, и обычно используется $m(\mathbf{x}) = 0$, так как гауссовские процессы гибкие и могут хорошо аппроксимировать функцию среднего. κ – это положительно определенная *ядерная* или *ковариационная функция*. Таким образом, гауссовский процесс – это распределение над функциями, форма которых (гладкость, ...) определяется \mathbf{K} . Если точки \mathbf{x}_i и \mathbf{x}_j считаются похожими ядром, значения функций в этих точках, $f(\mathbf{x}_i)$ и $f(\mathbf{x}_j)$, также должны быть похожими.

Пусть дан обучающий набор данных с функциональными значениями без шума \mathbf{f} на входах \mathbf{x} . Для прогнозирования \mathbf{f}_* на новых входах \mathbf{x}_* используется апостериорное распределение $p(\mathbf{f}_*|\mathbf{x}_*, \mathbf{x}, \mathbf{f})$. По определению гауссовского процесса совместное распределение наблюдений \mathbf{f} и предсказаний \mathbf{f}_* снова является гауссовским, что можно записать в виде

$$\begin{pmatrix} \mathbf{f} \\ \mathbf{f}_* \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^T & \mathbf{K}_{**} \end{pmatrix}\right)$$

где $\mathbf{K}_* = \kappa(\mathbf{x}, \mathbf{x}_*)$ и $\mathbf{K}_{**} = \kappa(\mathbf{x}_*, \mathbf{x}_*)$. Если известно N наблюдений (обучающих данных) и мы хотим сделать N_* предсказаний, тогда \mathbf{K} имеет размерность $N \times N$, матрица \mathbf{K}_* имеет размерность $N \times N_*$, а матрица \mathbf{K}_{**} имеет размерность $N_* \times N_*$. Стандартные преобразования позволяют записать искомое распределение следующим образом:

$$\begin{aligned} p(\mathbf{f}_* | \mathbf{x}_*, \mathbf{x}, \mathbf{f}) &= \mathcal{N}(\mathbf{f}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \\ \boldsymbol{\mu}_* &= \mathbf{K}_*^T \mathbf{K}^{-1} \mathbf{f} \\ \boldsymbol{\Sigma}_* &= \mathbf{K}_{**} - \mathbf{K}_*^T \mathbf{K}^{-1} \mathbf{K}_* \end{aligned}$$

Если же у нас есть обучающий набор данных, в котором у наблюдений присутствует шум $\mathbf{y} = \mathbf{f} + \boldsymbol{\varepsilon}$, где шум $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$ независимо добавляется к каждому наблюдению, то искомое распределение задается следующим образом:

$$p(\mathbf{f}_* | \mathbf{x}_*, \mathbf{x}, \mathbf{y}) = \mathcal{N}(\mathbf{f}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*) \quad (2)$$

$$\boldsymbol{\mu}_* = \mathbf{K}_*^T \mathbf{K}_y^{-1} \mathbf{y} \quad (3)$$

$$\boldsymbol{\Sigma}_* = \mathbf{K}_{**} - \mathbf{K}_*^T \mathbf{K}_y^{-1} \mathbf{K}_* \quad (4)$$

где $\mathbf{K}_y = \mathbf{K} + \sigma_y^2 \mathbf{I}$. Хотя уравнение (2) охватывает шум в обучающих данных, оно все еще представляет собой распределение над предсказаниями без шума \mathbf{f}_* . Чтобы также включить шум $\boldsymbol{\varepsilon}$ в предсказания \mathbf{y} , мы должны добавить σ_y^2 на диагональ матрицы $\boldsymbol{\Sigma}_*$

$$p(\mathbf{y}_* | \mathbf{x}_*, \mathbf{x}, \mathbf{y}) = \mathcal{N}(\mathbf{y}_* | \boldsymbol{\mu}_*, \boldsymbol{\Sigma}_* + \sigma_y^2 \mathbf{I})$$

используя определения $\boldsymbol{\mu}_*$ и $\boldsymbol{\Sigma}_*$ из уравнений (3) и (4) соответственно. Этих сведений достаточно для реализации гауссовых процессов и их применения к задачам регрессии.

3.2.3 Применение в наших целях

Итак, предлагается моделировать случайное трехмерное гауссовское поле и рассматривать его на поверхности цилиндра \mathcal{C} . Поскольку изображение дискретно, то цилиндр состоит из конечного числа точек. Например, если размеры изображения 50×100 , тогда \mathcal{C} есть объединение 50 слоев, каждый из которых представляет из себя 100 точек, расположенных в виде окружности.

Поскольку нам известен знак на полюсах, то предлагается потребовать (задать наблюдаемые значения в регрессии), чтобы поле принимало соответствующие значения вблизи полюсов (5 верхних и нижних окружностей на \mathcal{C}).

Пусть κ_1 и κ_2 – корреляционные функции. Тогда функция $\kappa(\mathbf{x}, \mathbf{x}') := \kappa_1(\mathbf{x}, \mathbf{x}') + \kappa_2(\mathbf{x}, \mathbf{x}')$ также является корреляционной функцией. Доказательство этого факта можно найти в [14]. На основе этого предлагается выбрать в нашем случае

$$\kappa_{\text{Periodic}}(\mathbf{x}, \mathbf{x}') = \exp \left(-2 \sum_i \frac{\sin^2 \left(\frac{\pi}{p} (x_i - x'_i) \right)}{\lambda} \right)$$

где p и λ – свободные параметры, отвечающие за длину периода и масштаб; и

$$\kappa_{\text{RBF}}(\mathbf{x}, \mathbf{x}') = \exp \left(-\frac{1}{2} (\mathbf{x} - \mathbf{x}')^\top \Theta^{-2} (\mathbf{x} - \mathbf{x}') \right)$$

где Θ – коэффициент масштаба.

Таким образом, мы используем нулевое среднее и следующую корреляционную функцию:

$$\kappa(\mathbf{x}, \mathbf{x}') = \alpha_1 \kappa_{\text{Periodic}}(\mathbf{x}, \mathbf{x}') + \alpha_2 \kappa_{\text{RBF}}(\mathbf{x}, \mathbf{x}')$$

Действительные коэффициенты α_i выбираются в процессе обучения алгоритма регрессии.

3.3 Примеры сгенерированных изображений

В результате удалось получить возможность генерировать бесконечное число изображений, удовлетворяющих нашим требованиям. На рис. 7 показаны некоторые примеры получившихся изображений.

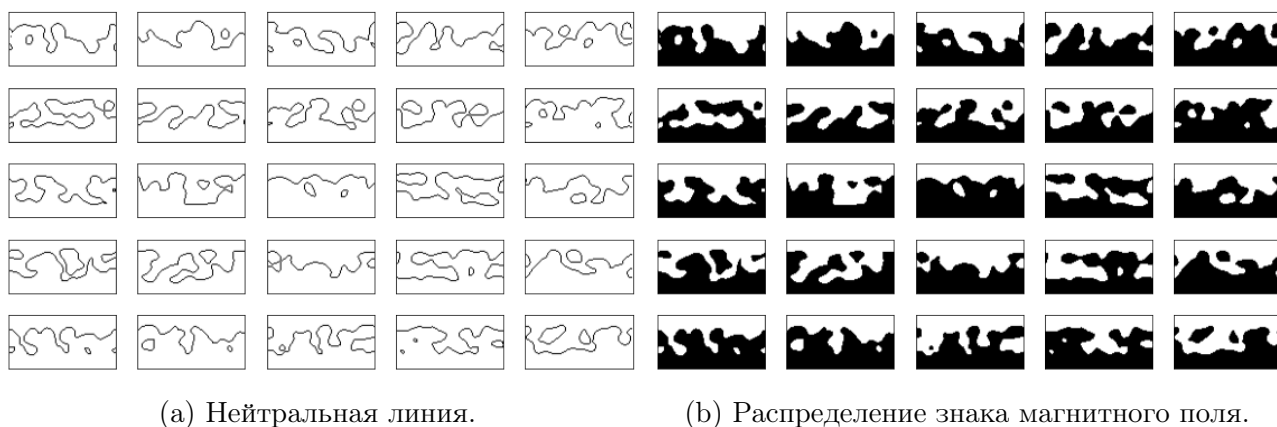


Рис. 7: Примеры изображений, которые можно получить предлагаемым нами методом. Размер каждого 50×100 .

Отметим некоторые важные практические преимущества нашего подхода:

- размер получаемых изображений может быть произвольным;
- природа гауссовского распределения такова, что важное условие смены знака при переходе через границу выполняется автоматически на получившихся изображениях;
- изменение параметров корреляционной функции κ позволяет гибко корректировать сложность структуры распределения знака магнитного поля и нейтральной линии в частности;
- благодаря фреймворку *GPyTorch* языка *python*, необходимые вычисления для работы алгоритма можно производить на GPU, что значительно ускоряет процесс обучения и вывода;
- в нашем случае мы моделируем трехмерное гауссовское поле на цилиндре, чтобы учесть условие «склеивания» краев. Вместо цилиндра, вложенного в трехмерное пространство, можно брать более сложные многообразия и не обязательно в \mathbb{R}^3 . Это позволит получать изображения, удовлетворяющие нетривиальным условиям, которые могут потребоваться в других областях.

4 Эксперимент на датасете

4.1 Структура эксперимента

Для проведения эксперимента были зафиксированы 100 изображений размера 50×100 . Затем с каждого из них удалялось по 0%, 50%, 65%, 80% и 90% информации о волокнах (см. рис. 8) случайным образом. Так получилось 5 наборов по 100 изображений, на каждом из которых и проводилось тестирование. Стоит отметить, что для каждого входного изображения производилось 5 запусков моделей. Это можно видеть на графиках из главы 4.3.

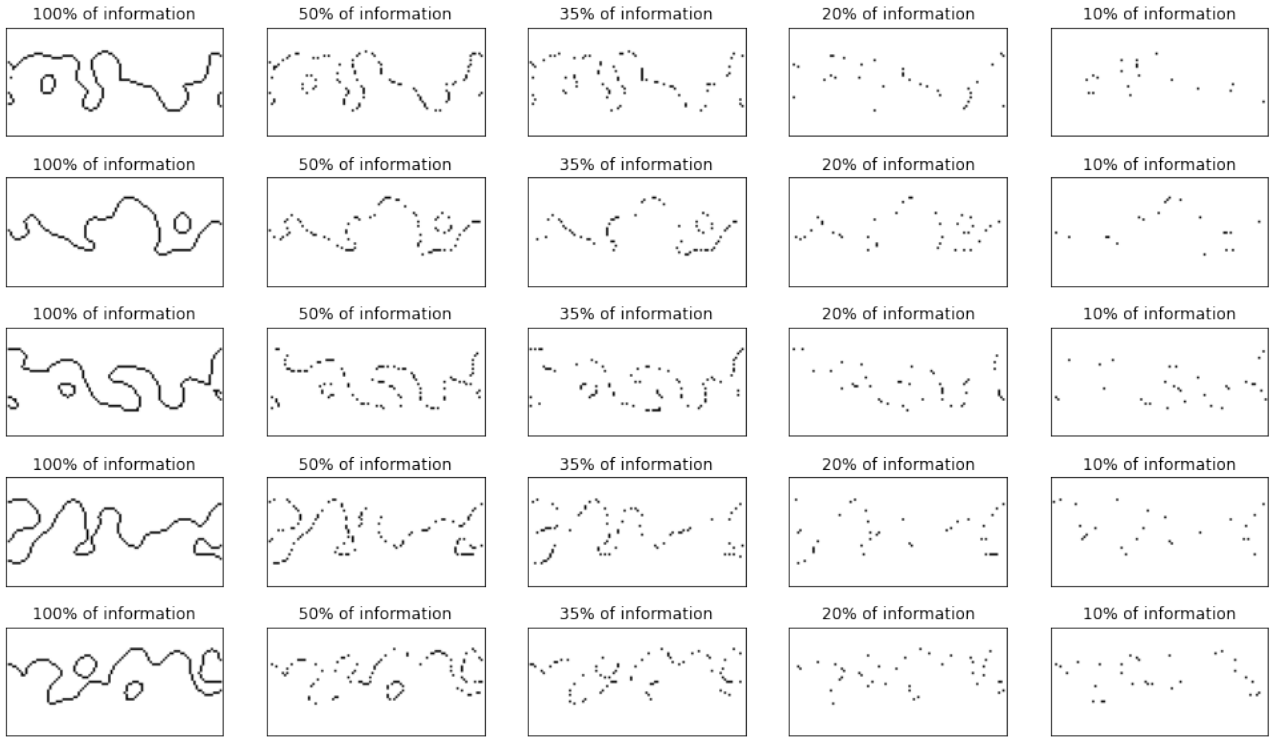


Рис. 8: Удаление информации о волокнах для 5 изображений (левый столбец).

4.2 Метрика качества

Каждое изображение (матрицу) размера $a \times b$ можно вытянуть в вектор длины $a \times b$. Пусть N – соответствующая длина вектора. Для каждого входного вектора \mathbf{x} мы имеем вектор \mathbf{y} , задающий истинное распределение знака. Тогда для предсказания $\hat{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ величина ошибки \mathcal{E} вычисляется следующим образом:

$$\mathcal{E}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

Можно считать, что наша функция f «раскрашивает» каждый пиксель. Тогда ошибку можно интерпретировать как *процент неверно раскрашенных пикселей*.

Важно понимать, что зная знак поля на полюсах, можно просто провести нейтральную линию в виде горизонтальной прямой по центру изображения и расставить знаки поля в соответствии со знаками на полюсах. В таком случае явно уже будет «угадано» не меньше половины распределения знака магнитного поля. Это объясняется спецификой задачи, однако такой подход, очевидно, не дает практической ценности.

Для «нормировки» нашей метрики был проведен дополнительный тест, на котором было оставлено лишь 5% информации о волокнах. Это можно считать псевдослучайным срабатыванием модели. На финальном графике на рис. 19 можно видеть, какое значение нашей метрики получается в этом случае.

4.3 Результаты

4.3.1 Полная информация о нейтральной линии

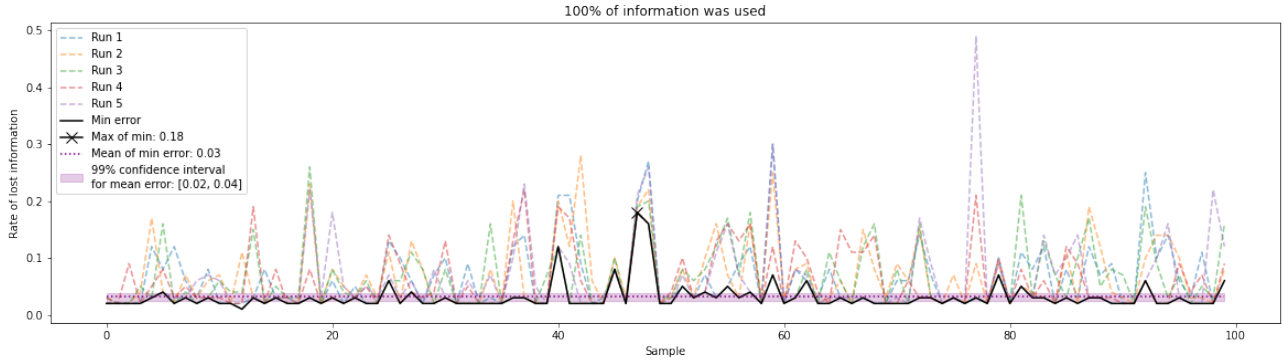


Рис. 9: Черная ломанная показывает наименьшую ошибку из 5 предсказаний (цветные ломанные) для каждого из 100 тестовых изображений.

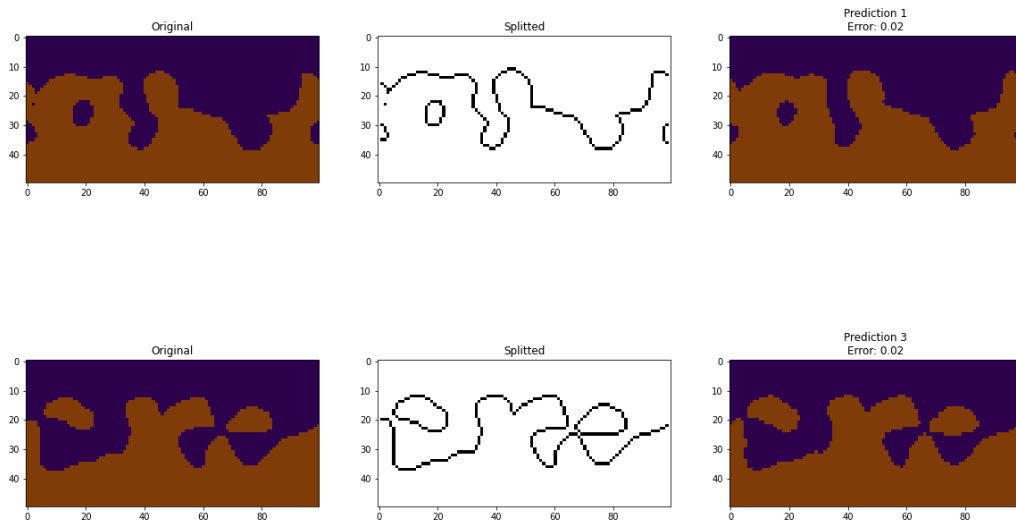


Рис. 10: Примеры предсказаний модели. Истинное распределение знака магнитного поля (слева), информация о нейтральной линии, входное изображение (по центру), предсказание модели (справа).

При полном доступе к информации о положении нейтральной линии ошибка предсказаний составила $\mathcal{E} = 3 \pm 1\%$. Модель восстанавливает структуру магнитного поля на **97%**.

Полученное значение *подтверждают состоятельность подхода*. Для наглядности на рис. 10 представлены примеры предсказаний модели.

4.3.2 50% информации о нейтральной линии

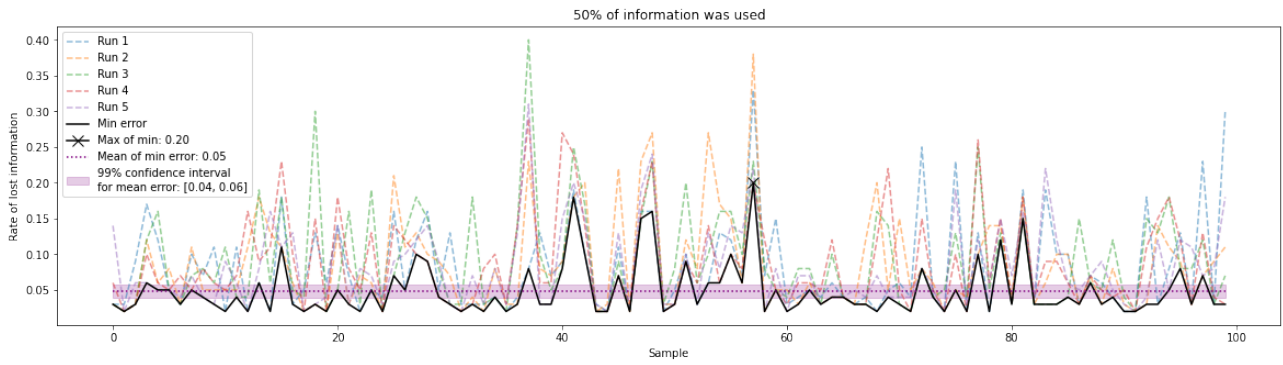


Рис. 11: Черная ломанная показывает наименьшую ошибку из 5 предсказаний (цветные ломанные) для каждого из 100 тестовых изображений.

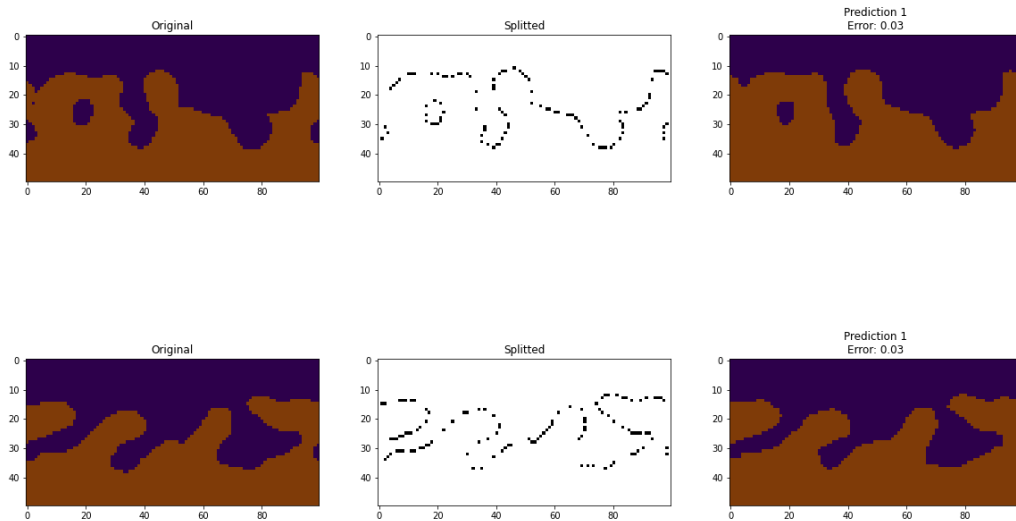


Рис. 12: Примеры предсказаний модели. Истинное распределение знака магнитного поля (слева), информация о нейтральной линии, входное изображение (по центру), предсказание модели (справа).

Имея доступ к 50% информации о положении нейтральной линии ошибка предсказаний составила $\mathcal{E} = 5 \pm 1\%$. Модель восстанавливает структуру магнитного поля на **95%**. Для наглядности на рис. 12 представлены примеры предсказаний модели.

4.3.3 35% информации о нейтральной линии

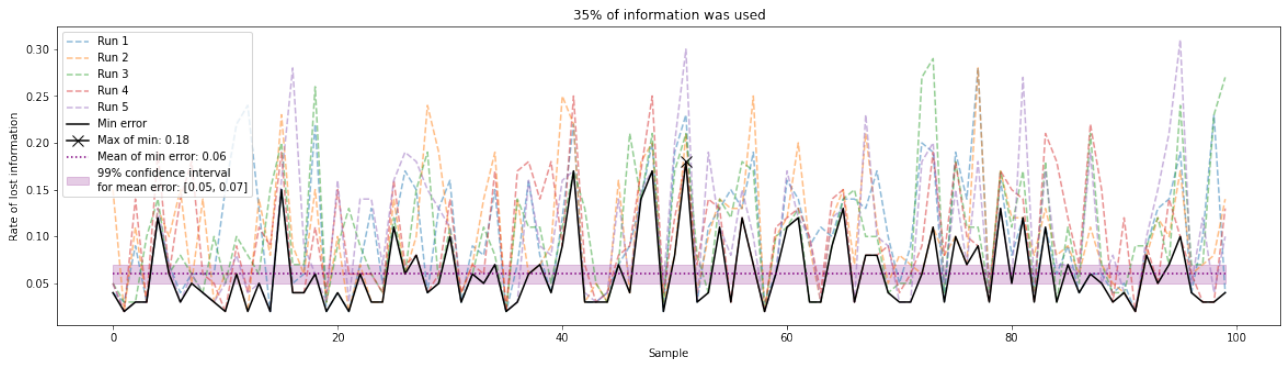


Рис. 13: Черная ломанная показывает наименьшую ошибку из 5 предсказаний (цветные ломанные) для каждого из 100 тестовых изображений.

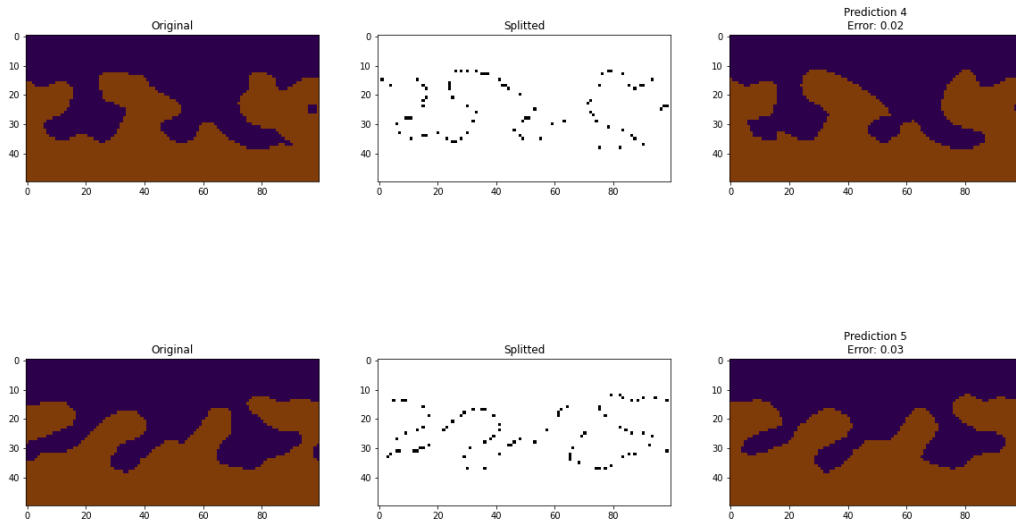


Рис. 14: Примеры предсказаний модели. Истинное распределение знака магнитного поля (слева), информация о нейтральной линии, входное изображение (по центру), предсказание модели (справа).

Имея доступ к 35% информации о положении нейтральной линии ошибка предсказаний составила $\mathcal{E} = 6 \pm 1\%$. Модель восстанавливает структуру магнитного поля на **94%**. Для наглядности на рис. 14 представлены примеры предсказаний модели.

4.3.4 20% информации о нейтральной линии

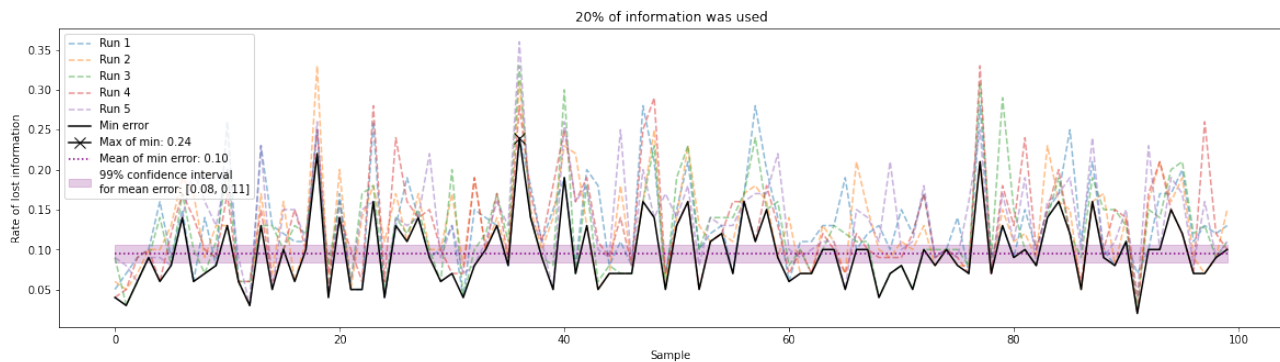


Рис. 15: Черная ломанная показывает наименьшую ошибку из 5 предсказаний (цветные ломанные) для каждого из 100 тестовых изображений.

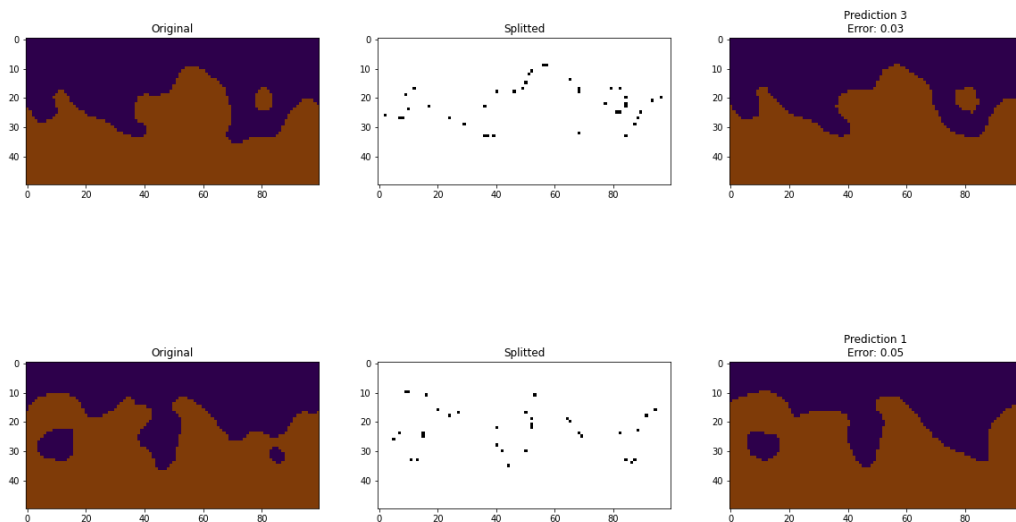


Рис. 16: Примеры предсказаний модели. Истинное распределение знака магнитного поля (слева), информация о нейтральной линии, входное изображение (по центру), предсказание модели (справа).

Имея доступ к 20% информации о положении нейтральной линии ошибка предсказаний составила $\mathcal{E} = 9.5 \pm 1\%$. Модель восстанавливает структуру магнитного поля на **90%**. Для наглядности на рис. 16 представлены примеры предсказаний модели.

4.3.5 10% информации о нейтральной линии

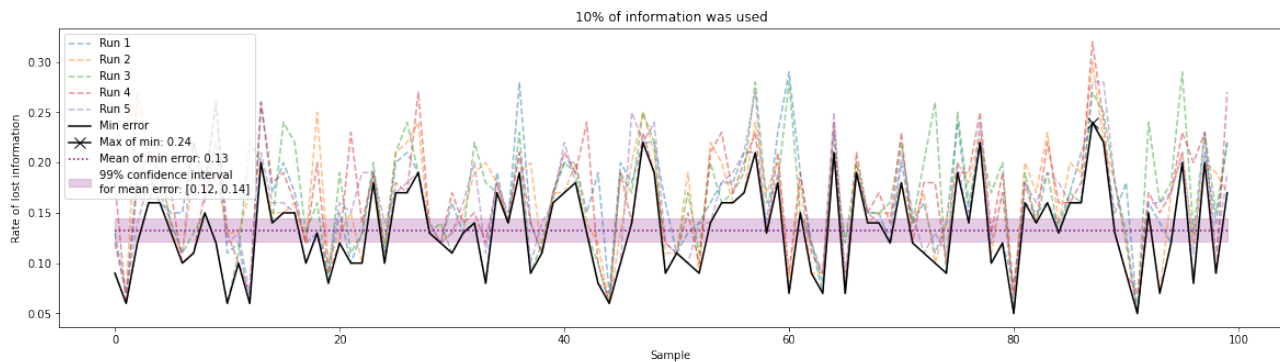


Рис. 17: Черная ломанная показывает наименьшую ошибку из 5 предсказаний (цветные ломанные) для каждого из 100 тестовых изображений.

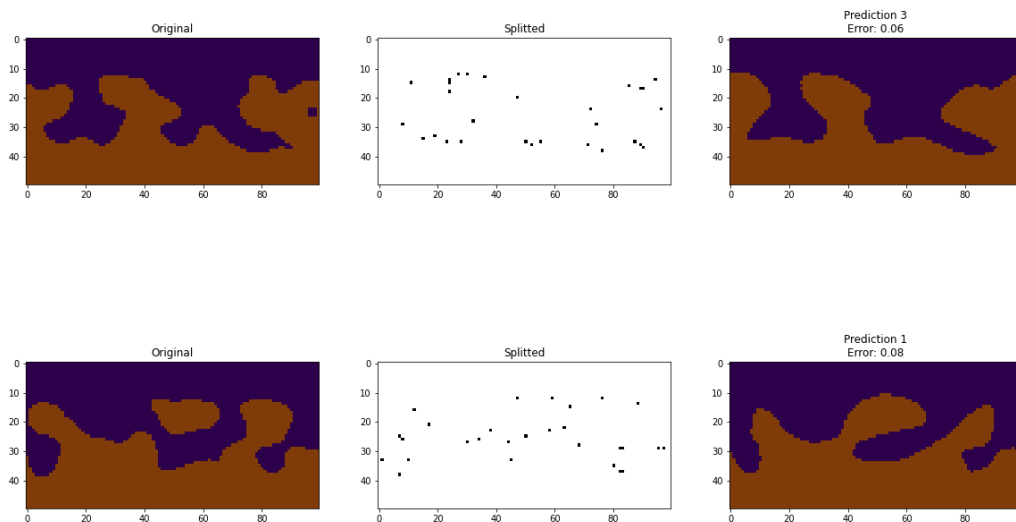


Рис. 18: Примеры предсказаний модели. Истинное распределение знака магнитного поля (слева), информация о нейтральной линии, входное изображение (по центру), предсказание модели (справа).

Имея доступ к 10% информации о положении нейтральной линии ошибка предсказаний составила $\mathcal{E} = 13 \pm 1\%$. Модель восстанавливает структуру магнитного поля на **87%**. Для наглядности на рис. 18 представлены примеры предсказаний модели.

4.4 Выводы из эксперимента

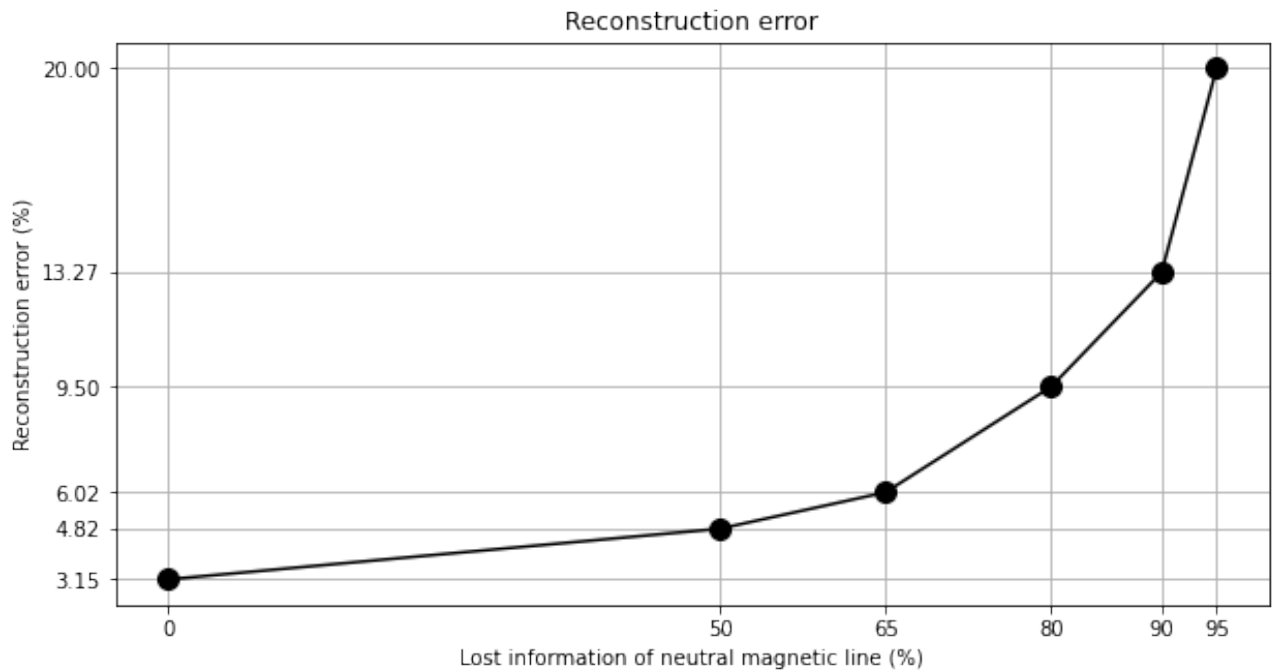


Рис. 19: График, показывающий рост ошибки предсказаний \mathcal{E} в зависимости от процента скрытой информации об истинном положении нейтральной линии.

Как упоминалось ранее, был проведен дополнительный эксперимент (крайняя правая точка на графике на рис. 19), показывающий величину ошибки при почти что полном отсутствии информации.

Полученные значения демонстрируют способность модели хорошо восстанавливать изображения, на которых отсутствует вплоть до 90% информации об истинном положении нейтральной линии. Это подчеркивает потенциал предложенного подхода для анализа наблюдательных данных по солнечной активности.

По ссылке [16] находится программный код и все полученные предсказания.

5 Эксперимент на реальных картах

После успешного тестирования модели на сгенерированных синтетических изображениях, возник интерес к тому, как поведет себя модель на реальных данных о волокнах (как например на рис. 2).

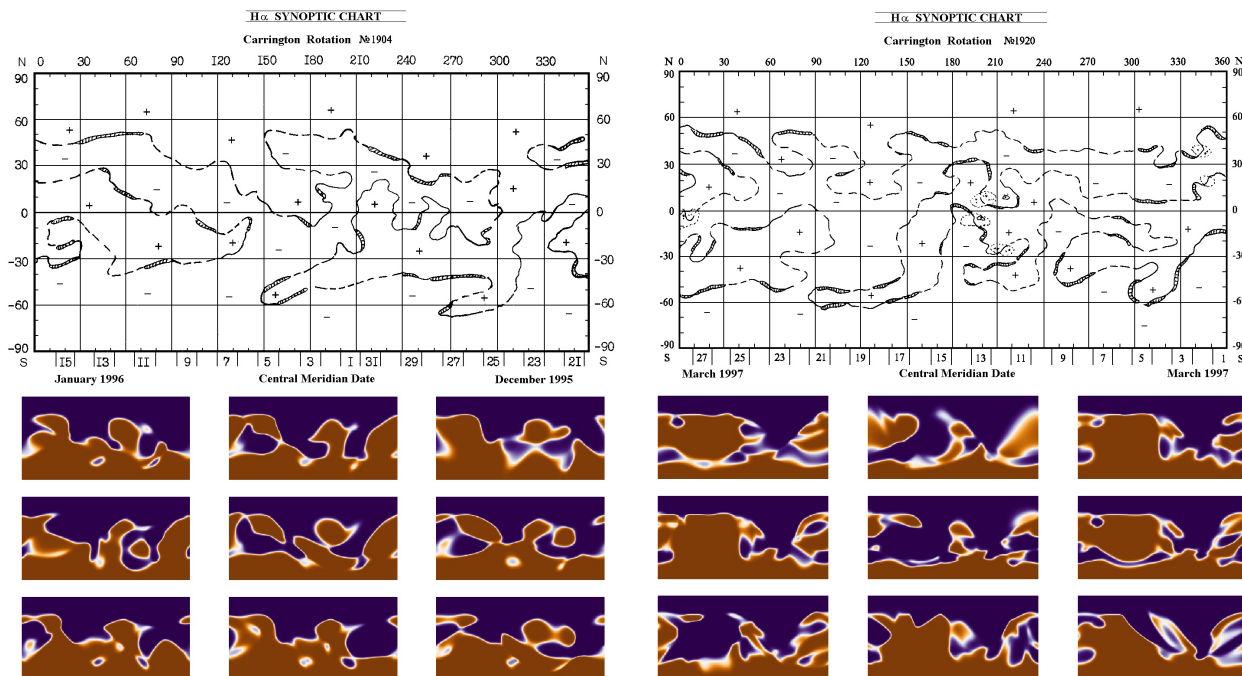


Рис. 20: Сравнение предсказаний модели и синоптических карт, составленных астрономами-наблюдателями. Для каждой карты было произведено 9 запусков модели.

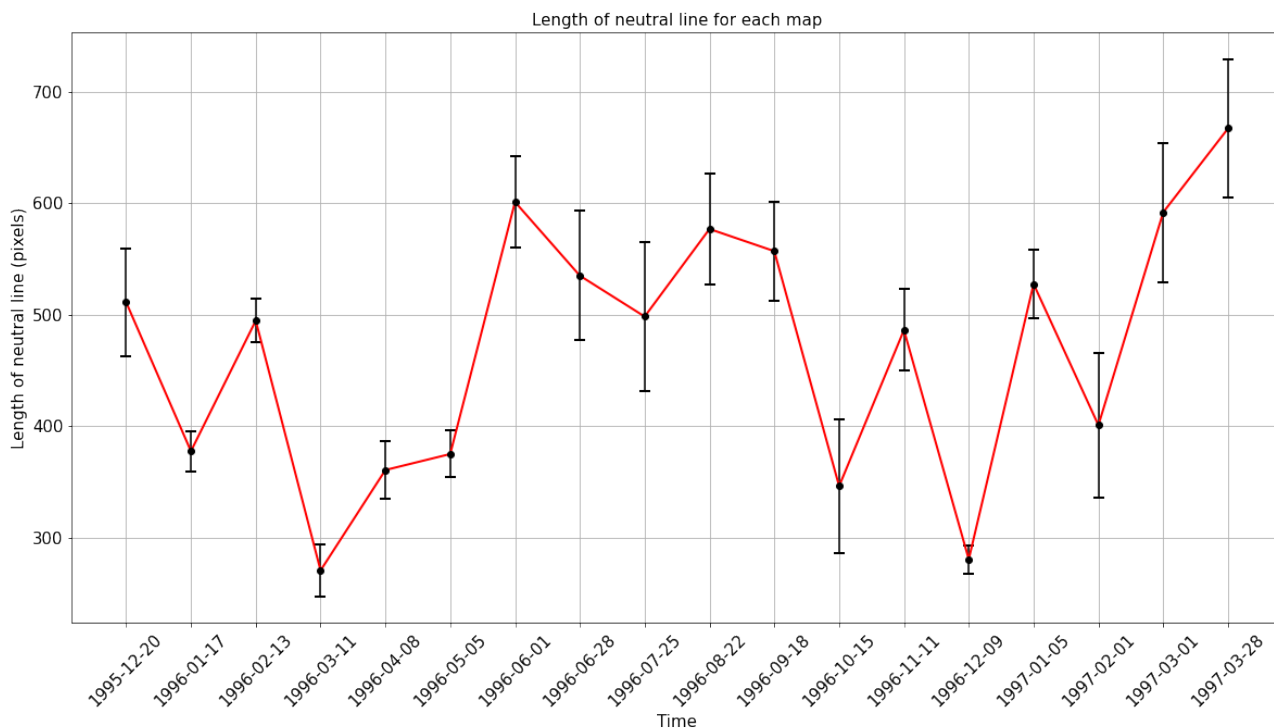


Рис. 21: Длина нейтральной линии на предсказаниях. На оси абсцисс отмечены даты, соответствующие картам с информацией о волокнах за этот период.

Стоит отметить, что оценка предсказаний нашего алгоритма является нетривиальной задачей в этом случае. Синоптические карты представляют собой изображение, в то время как карты с волокнами заданы текстовыми файлами с координатами пикселей, соответствующих наблюдаемым волокнам. Поэтому, для сравнения наших предсказаний с синоптическими картами требуется разработка дополнительного алгоритма, который мог бы представить синоптическую карту (изображение) в виде матрицы без лишней информации.

Вместо этого было решено исследовать динамику изменения длины нейтральной линии на предсказаниях. Из получившегося графика на рис. 21 можно сделать два важных вывода.

1. Небольшая величина разброса длины средней линии свидетельствует об устойчивости нашего алгоритма.
2. Полученные значения не выглядят случайно. Это может свидетельствовать о некоей закономерности в данных, для которых строились предсказания. В главе 6.1 дадим возможное объяснение этому явлению.

6 Направления дальнейших исследований

6.1 Последовательные карты

Если сравнить две или несколько последовательных синоптических карт, то зачастую можно заметить схожесть в распределении знака магнитного поля Солнца на них. Это связано со скоростью и непрерывностью изменений магнитных полей.

В дальнейших исследованиях стоит рассмотреть возможность использования информации, содержащейся в нескольких последовательных синоптических картах. Описываемый в этой работе подход основывается на одной карте (изображении), края которого склеиваются путем наложения на цилиндр. В частности, можно склеивать соответственные края двух или нескольких последовательных синоптических карт для получения более детальной информации о распределении знаков магнитного поля Солнца сразу за несколько оборотов вокруг своей оси. Это может улучшить точность результатов нашего подхода.

Стоит отметить, что функционал используемой модели позволяет производить процесс обучения на нескольких картах. Исследования показали, что в некоторых случаях описываемый в этой главе подход может давать прирост в качестве предсказаний. Например, были попытки обучить модель сразу на двух картах, помещенных на два концентрических цилиндра.

6.2 Многомасштабный подход

Идея подхода заключается в следующем: вначале восстанавливается глобальная структура магнитного поля, а затем происходит уточнение деталей. Обратимся к рис. 22. На нем видно, как по входным изображениям (второй столбец) строятся карты, отражающие глобальную структуру поля (крайний правый столбец). Для наглядности в левом столбце приведены оригиналы изображений.

Предлагается получать предсказания последовательно: сначала обучить модель для правого изображения, затем для соседнего слева и так дойти до входного. Таким образом происходит постепенное уточнение деталей.

Эксперименты показали, что такой подход в некоторых случаях дает не только выигрыш в качестве предсказаний, но и во времени обучения⁵.

⁵Входные изображения на рис. 22 имеют размер 50×100 , а изображения из правого столбца – 5×10 .

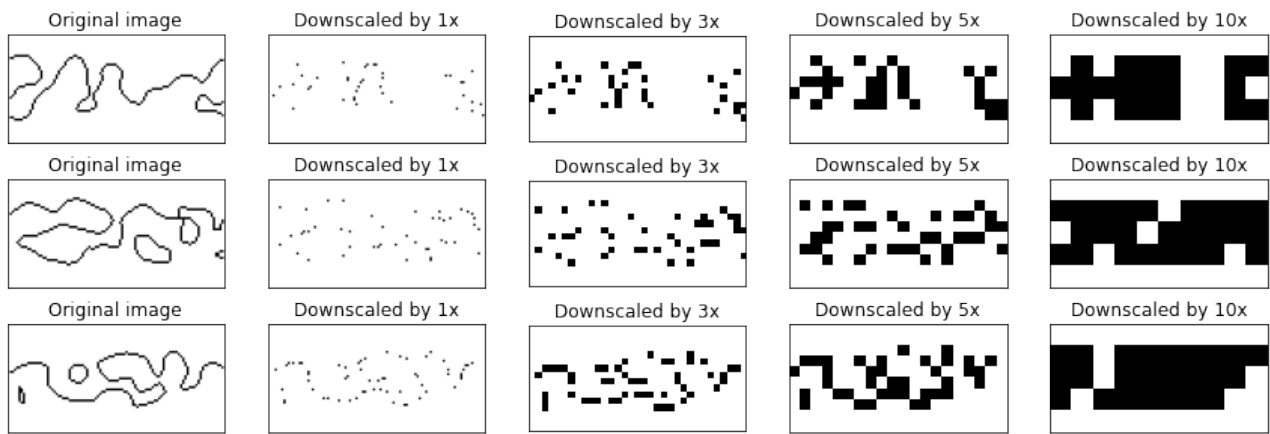


Рис. 22: Пример последовательности карт, получаемых при многомасштабном подходе.

7 Выводы

- Предложена модель, основанная на машинном обучении, способная восстанавливать нейтральную линию.
- Получены оценки точности предсказаний модели на синтетических данных.
- Дополнительно предложена методика создания датасета синтетических изображений.
- Модель апробирована на реальных данных.

8 Заключение

- Результаты тестирования подчеркивают потенциал предложенного подхода для анализа солнечных данных.
- В перспективе мы рассчитываем, что с помощью предложенного метода будет проведена обработка большого массива наблюдательных данных, что позволит создать новый однородный ряд данных, представляющий интерес для изучения в контексте физики Солнца.
- Методика создания синтетического датасета позволяет генерировать изображения, вообще говоря, произвольной сложности, что может быть полезно и в задачах вне контекста физики Солнца.
- Результаты данной работы были представлены на конференции «Ломоносов» [15].
- По ссылке [16] находится репозиторий с полным программным кодом и результатами всех экспериментов, описанных в данной работе.

Список литературы

- [1] Садыков В. М., Илларионов Е. А. Как машинное обучение помогает изучать Солнце. Земля и вселенная. — 2021. — № 4. — С. 35–45.
- [2] Solar large-scale magnetic field and cycle patterns in solar dynamo / V. N. Obridko, V. V. Pipin, D. D. Sokoloff, A. S. Shibalova. Monthly Notices of the Royal Astronomical Society. — 2021.
- [3] McIntosh P. S. Solar magnetic fields derived from hydrogen alpha filtergrams. Reviews of Geophysics. — 1972. — Т. 10. — №. 3. — С. 837-846.
- [4] Ipson S. S. et al. Automated technique for comparison of magnetic field inversion lines with filament skeletons from the Solar Feature Catalogue. Solar Physics. — 2005. — Т. 228. — С. 399-421.
- [5] Makarenko N. G., Karimova L. M., Novak M. M. Investigation of global solar magnetic field by computational topology methods. Physica A: Statistical Mechanics and its Applications. — 2007. — Т. 380. — С. 98-108.
- [6] Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators. Neural networks. — 1989. — Т. 2. — №. 5. — С. 359-366.
- [7] Cunningham P., Cord M., Delany S. J. Supervised learning. Machine learning techniques for multimedia: case studies on organization and retrieval. — 2008. — С. 21-49.
- [8] Ch R., MAP M.. L.. Bayesian learning. book: Machine Learning. McGraw-Hill Science/Engineering/Math. — 1997. — С. 154-200.
- [9] Nadaraya E. A. On estimating regression. Theory of Probability & Its Applications. — 1964. — Т. 9. — №. 1. — С. 141-142.
- [10] Watson G. S. Smooth regression analysis. Sankhyā: The Indian Journal of Statistics, Series A. — 1964. — С. 359-372.
- [11] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21-27.
- [12] Rasmussen, Carl E., and Christopher K. I. Williams. Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning series, MIT Press, 2006.
- [13] Булинский А. В., Ширяев А. Н. Теория случайных процессов. — ФИЗМАТЛИТ Москва, 2005. — 400 с
- [14] Gneiting T. Compactly supported correlation functions. Journal of Multivariate Analysis. — 2002. — Т. 83. — №. 2. — С. 493-508.
- [15] Киселюс В. Восстановление нейтральных линий магнитного поля Солнца методами машинного обучения. XXX Международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов». — 2023. <https://istina.msu.ru/conferences/presentations/550922838/>
- [16] <https://github.com/V-Kisielius/Diploma>