Сравнительный анализ методов выявления мошеннических транзакций

Igor A. Kondratev1, Vadim V. Bazanov2, Daniil A. Uskov3, Anna V. Kuchebo4  
Department of Computer Systems and Technologies  
National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)Moscow, Russian Federation  
1Igor.kondratev17@yandex.ru, 2vadimbaz7@gmail.com, 3daniiluskov337@gmail.com, 4a.kuchebo@mail.ru Tatyana E. Sereda  
Department of Automated Control System  
Moscow Automobile and Road Construction State Technical University (MADI)  
Moscow, Russian Federation  
sereda\_tasya@mail.ru

*Аннотация -* Данная статья посвящена сравнительному анализу моделей машинного обучения, используемых в целях идентификации мошеннических транзакций. В ходе работы были рассмотрены модели трёх различных алгоритмов, а также произведена их оптимизация под поставленную задачу. Было осуществлено сравнение полученных результатов работы моделей, выявлены достоинства и недостатки для каждой, приведены рекомендации по их использованию, сделаны выводы. Эксперимент проводился с целью оценки эффективности различных моделей машинного обучения в задачах классификации транзакций.

*Ключевые слова – Нейронная сеть, машинное обучение, банковские транзакции, мошенничество с кредитными картами.*

# Introducion

С каждым годом увеличивается количество людей, пользующихся безналичной формой оплаты и использующих банковские карты в виде оплаты через интернет. К сожалению, как и любой другой сфере, связанной с финансами, организации и пользователи интернет приложений при работе с банковскими картами сталкиваются с мошенничеством.

Согласно опросу многонационально сетевого сервиса PWC[1] на 2018 было 49% компаний из 7200 опрошенных сталкивались с мошенничеством, этот показатель вырос на 13% по сравнению с тем же опросом на 2016 год. Мошенничество растет с каждым годом и развивается напрямую с развитием информационных технологий. За последние несколько лет заметна тенденция роста числа мошенничеств в этой области, несмотря на непрерывное совершенствование систем защиты от них. На данный момент, мошенничество чаще наблюдается в уязвимости систем защиты банковских карт и систем осуществления транзакций[2].

В наше время численность мошеннических операций интенсивно растет. Мошенничество с банковскими картами широко распространено в таких европейских странах, как Россия, Англия, Франция, Германия и Испания.

Большинству людей и организаций, участвующих в банковских операциях и попавших под влияние мошенников, грозят значительные материальные потери, а также увеличение репутационных рисков для организаций. В данном случае, e-commerce отрасль не является исключением, так как мошеннические операции повлекут за собой недоверие и негативное отношение пользователей интернета к данной области, что в свою очередь приведет к снижению спроса на электронные платежи[3].

В машинном обучении такие проблемы, как обнаружение мошенничества, обычно называют проблемами классификации – предсказанием одного из двух возможных значений на основе данных наблюдений. Когда речь заходит об обнаружении мошенничества с платежами по кредитным картам, задача классификации состоит в создании моделей с достаточным интеллектом, чтобы правильно классифицировать транзакции как законные или мошеннические, основываясь на данных транзакции, таких как сумма, продавец, местоположение, время и другие[4].

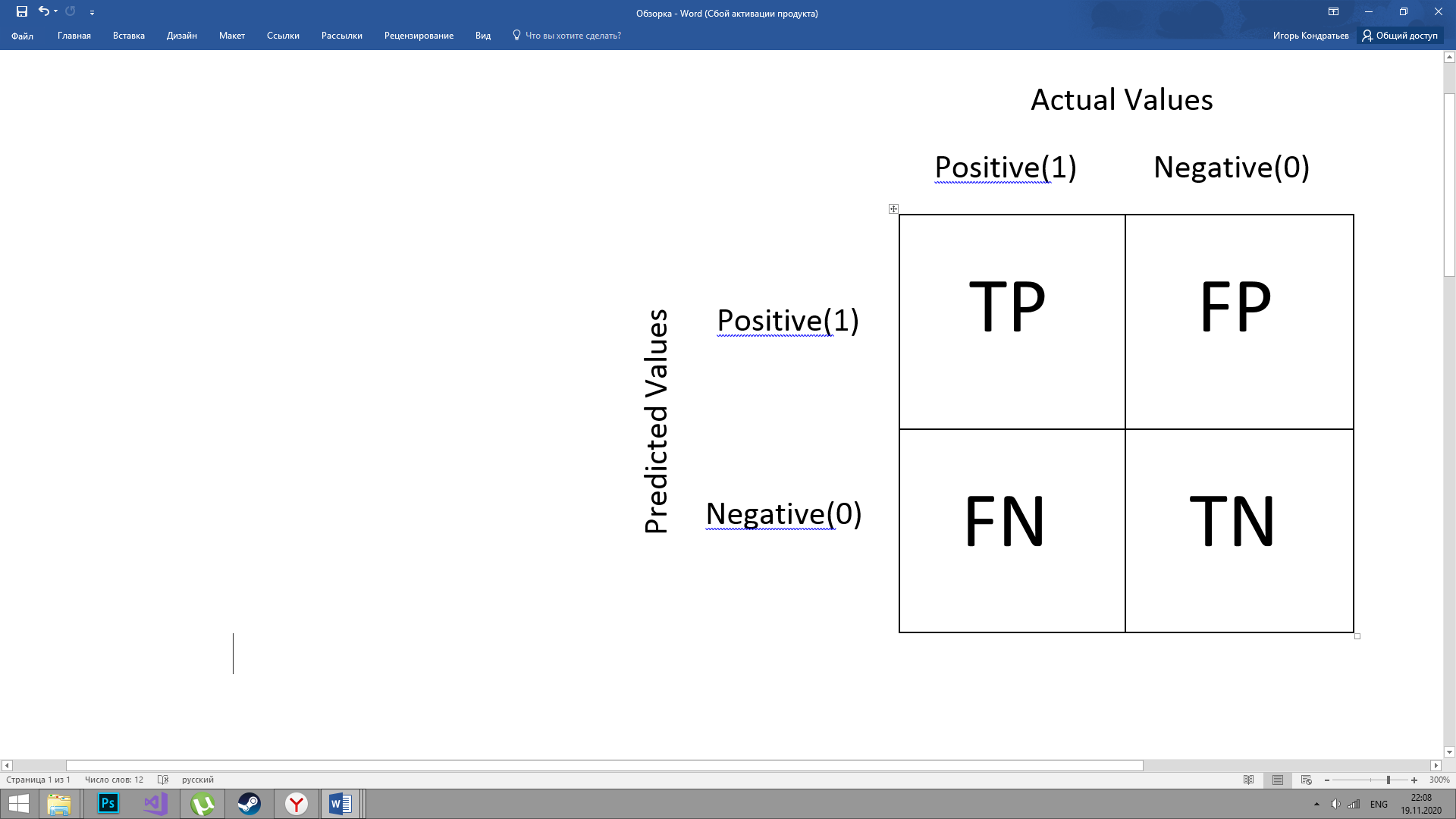
Финансовые махинации в сфере электронных банковских переводов по-прежнему составляют значительные суммы денежных средств. Мошенники по всему миру постоянно ищут новые способы совершения различного рода хищений[5]. Полагаясь исключительно на основанные на правилах запрограммированные системы для обнаружения финансового мошенничества, невозможно обеспечить надлежащую безопасность банковских переводов. Современные системы обнаружения мошенничества способны предотвращать одну двенадцатую всех обрабатываемых транзакций, что приводит к потерям в миллиарды долларов[6]. Именно здесь машинное обучение является уникальным решением для такого рода проблем.

Существует множество решений для защиты от мошенничества, которые основаны на алгоритмах машинного обучения[7], также часто используются стандартные способы защиты, основанные на сопоставлении данных перевода с определённым набором правил, списков и фильтров. Но в большинстве случаев модели на основе алгоритмов машинного обучения показывают большую точность. В современных системах часто используется такие алгоритмы, как случайный лес или байесовский классификатор, но также присутствуют решения использующие нейронные сети вместо классических алгоритмов[8]. Но модели глубокого обучения требуют гораздо большее количество ресурсов и времени для их использования. А так как для современных систем скорость работы является приоритетным аспектом в данной работе проводится сравнение работы алгоритмов байесовского классификатора и двух нейросетей для объективной оценки их эффективности.

Практическая значимость: проводимое сравнение выявит эффективность использования технологий глубокого обучения в задачах классификации банковских транзакций, и поможет провести объективную оценку целесообразности использования тех или иных алгоритмов машинного обучения в проектах связанных с выявлением мошенничеств.

# Science matireals and methods

В качестве меры сравнения моделей будет использоваться матрица ошибок, образец которой представлен на Fig. 1:



1. Образец матрицы ошибок

* TP - истино-положительное решение
* TN - истино-отрицательное решение
* FP - ложно-положительное решение
* FN - ложно-отрицательное решение

Из матрицы ошибок можно выявить такие характеристики моделей, как полнота, точность и F-мера.

Полнота Recall:

(1)

Из всех положительных классов, сколько мы предсказали правильно.

Точность Precision:

(2)

Из всех классов, сколько мы предсказали правильно.

F-мера:

(3)

F-мера это оптимальная характеристика модели, сочетающая в себе точность и полноту модели. В этой формуле полнота и точность имеют одинаковый вес, поэтому если одна из характеристик будет падать F-мера модели также будет уменьшаться.

Как уже говорилось ранее, в данной работе будут сравниваться 3 модели машинного обучения:

* Naive bayes classifier
* Нейронная сеть прямого распространения
* Автокодировщик

Подробнее о каждой из них:

* Наивный байесовский классификатор.

Этот алгоритм работает на основе формулы Байеса.

Формула Байеса:

P(A | B) = {\displaystyle P(A\mid B)={\frac {P(B\mid A)\,P(A)}{P(B)}}} (5)

{\displaystyle P(A)} P(A)  — априорная вероятность гипотезы *A*

P(A | B) {\displaystyle P(A\mid B)} — вероятность гипотезы *A* при наступлении события *B* (апостериорная вероятность);

P(B | A) {\displaystyle P(B\mid A)} — вероятность наступления события *B* при истинности гипотезы *A*;

P(B) {\displaystyle P(B)} — полная вероятность наступления события *B*

Цель классификации-понять, к какому классу относится объект, поэтому нам нужна не сама вероятность, а наиболее вероятный класс. Байесовский классификатор использует оценку апостериорного максимума для определения наиболее вероятного класса.

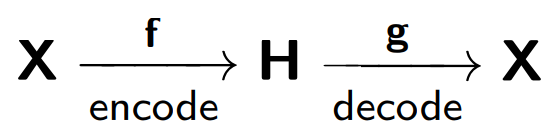
Этот алгоритм часто используется в системах обнаружения мошеннических переводов, а также обладает высокой степенью точности и скоростью выполнения, именно поэтому он был выбран для сравнения[9].

* Нейронная сеть прямого распространения.

Данный тип нейронных сетей часто используется для задач классификации, что нам и требовалась в нашей задаче.

Нейронная сеть прямого распространения-это классический тип нейронной сети, в которой связи между нейронами не образуют замкнутых циклов. Нейронные сети прямого распределения являются первым и наиболее простым типом для проектирования классических нейронных сетей. В этой сети информация распространяется только в одном направлении (вперед) от входов через скрытые слои (если таковые имеются) к выходам.

* Автокодировщик

Формально автокодировщик это нейронная сеть, которая обучается восстанавливать объекты, принятые сетью в качестве входных данных. 

1. Схема устройства автокодировщика

Автоэнкодер состоит из двух частей, как показано на Fig.2: кодера f, который кодирует выборку X в его внутреннее представление H, и декодера g, который восстанавливает исходную выборку. Таким образом, автокодер пытается сопоставить восстановленную версию каждого образца объекта с исходным объектом.

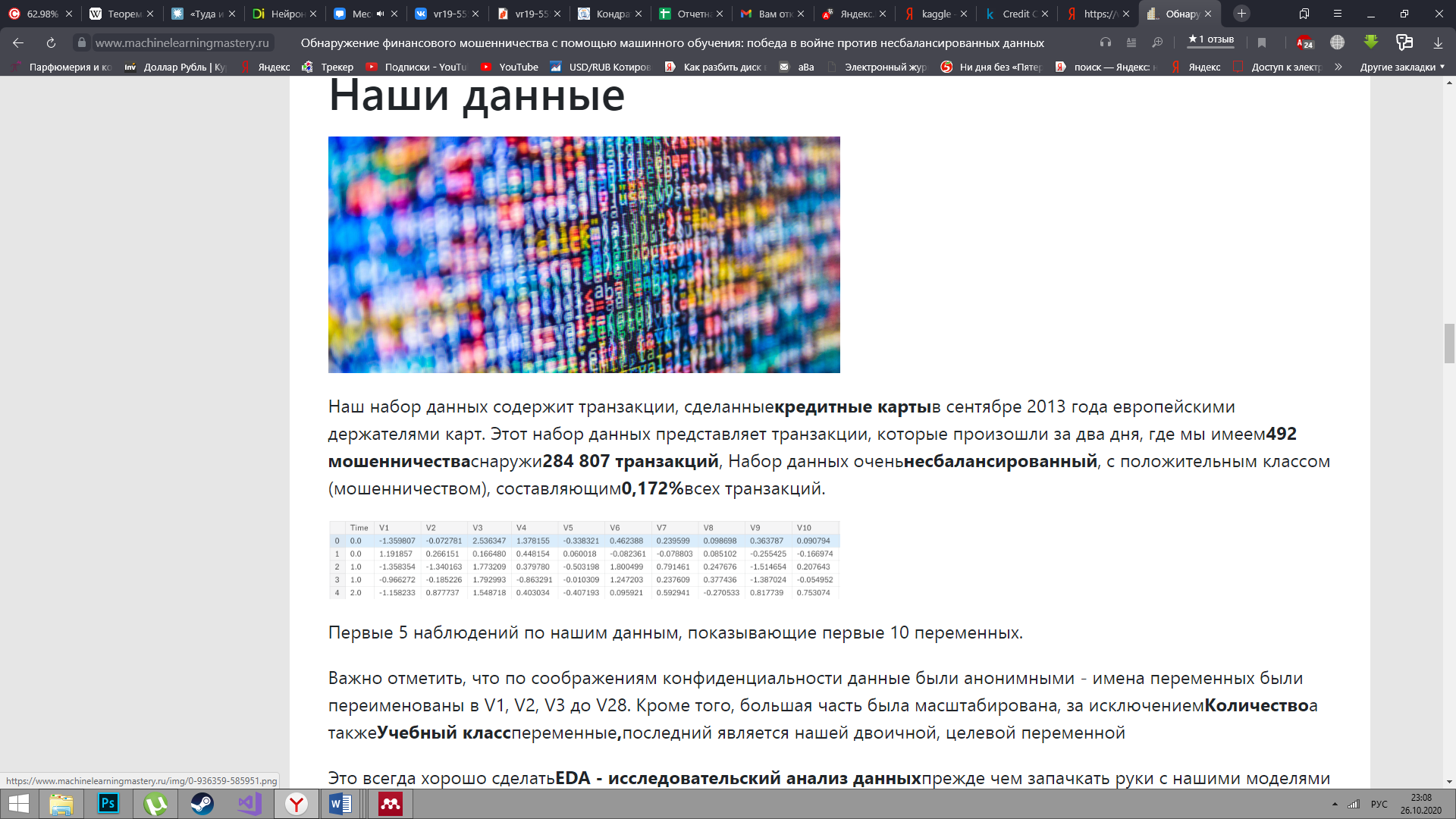
При обучении автокодировщика минимизируется следующая функция:

(4)

Где за **r(X)** обозначается восстановленная версия исходного объекта, а за **X** – сам объект.

Данная модель нейронной сети может быть также крайне эффективна в задаче выявления мошеннических банковских переводов, что было доказано в статье[10].

На вход каждой модели подавался один и тот же датасет, часть которого показана на Fig.3:



1. Образец используемого датасета

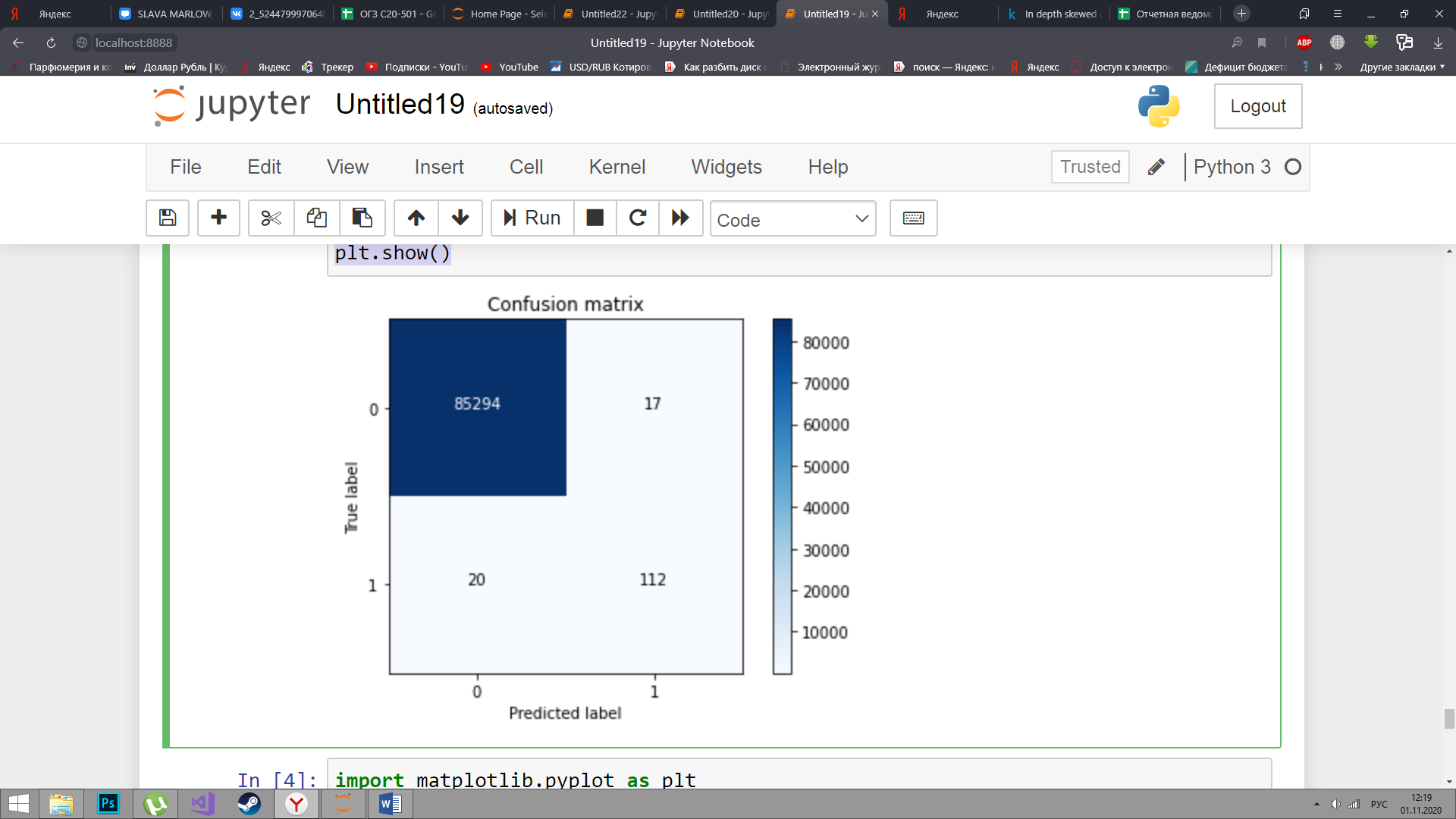
Данный датсет содержит в себе данные о 284,807 совершенных транзакциях, содержит мошеннические и не мошеннические транзакции, включает в себя столбцы с данными, где известны только время и сумма платежа, остальные данные помечены наименованиями V1..Vn (конфиденциальные данные). Единственное, что мы знаем, это то, что неизвестные столбцы уже масштабированы, т.е. данные прошли преобразование PCA(principal component analysis) – метод главных компонент, кроме времени и суммы. Среднее значение всех сделанных платежей составляет приблизительно 88 долларов. Большинство транзакций были без мошенничества (99,83%), в то время как мошенничество происходит (0,17%) в данном наборе данных.

Алгоритм проведения эксперимента:

* Разделение датасета на обучающую и тестовую части
* Обучение моделей и их оптимизация
* Обработка результатов работы моделей на тестовой части датасета
* Построение матриц ошибок, сравнение моделей, выводы

# Results

Нейронная сеть прямого распространения:



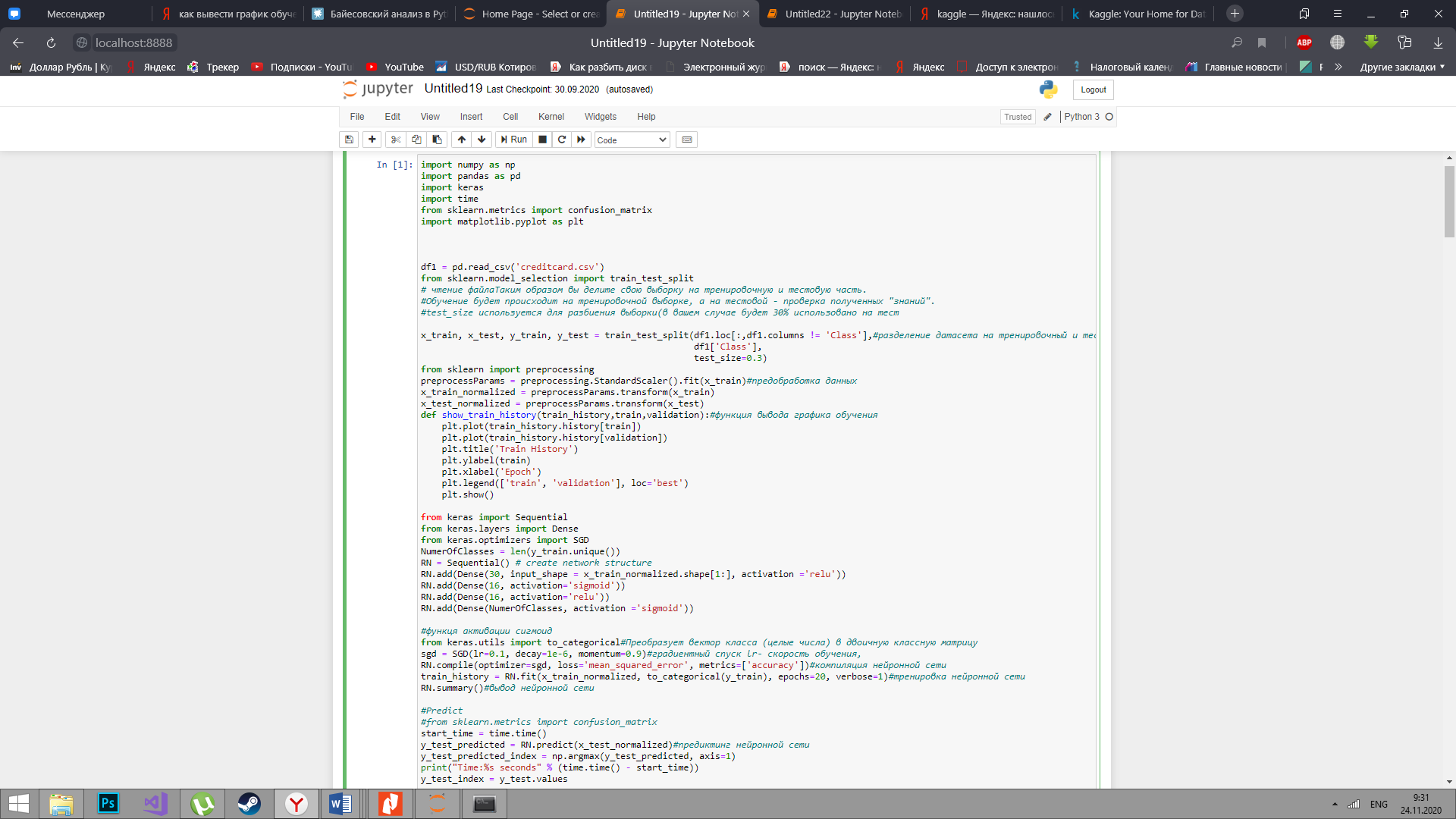
1. Матрица ошибок нейронной сети прямого распространения

* Среднее время оценки транзакции: 9,68189 мкс
* Точность на тестовом наборе: 99.9%
* F-мера: 0.999783
* Характеристики модели:
* Нейросеть состоит из 4 слоёв:

1 слой(входной): 30 нейронов, функция активации - relu

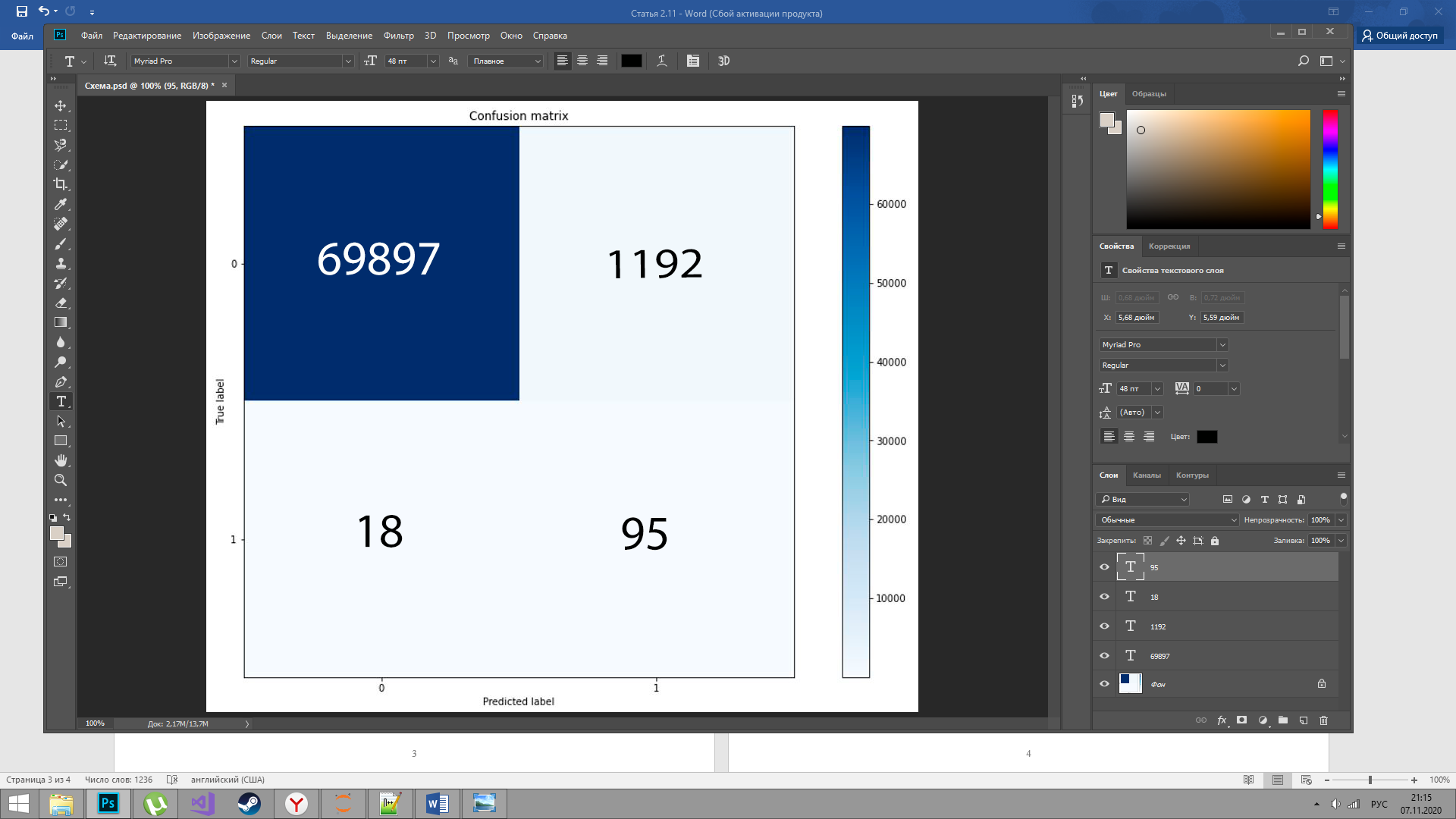
2,3 слои(скрытые): 16 нейронов, функции активации - sigmoid, relu

4 слой(выходной): 2 нейрона, функция активации – sigmoid

Часть кода с построением модели показана на Fig.5.

1. Построение модели нейронной сети прямого распространения

Автокодировщик:



1. Матрица ошибок автокодировщика

* Среднее время оценки транзакции: 14,4049 мкс
* Точность на тестовом наборе: 98.3%
* F-мера: 0.991419
* Характеристики модели:
* Нейросеть состоит из 7 слоёв:

1 слой(входной) ­­– 30 нейронов

2 слой(encoder) – 18 нейронов, функция активации tanh

3 слой(encoder) – 10 нейронов, функция активации elu

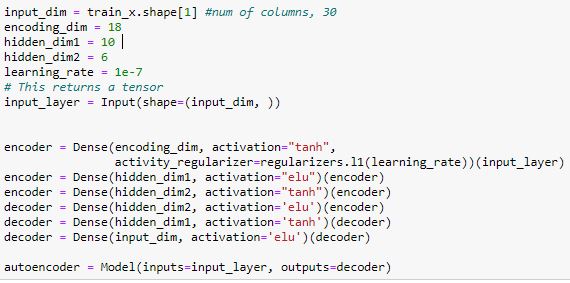
4 слой(encoder) – 6 нейронов, функция активации tanh

5 слой(decoder) – 6 нейронов, функция активации elu

6 слой(decoder) – 10 нейронов, функция активации tanh

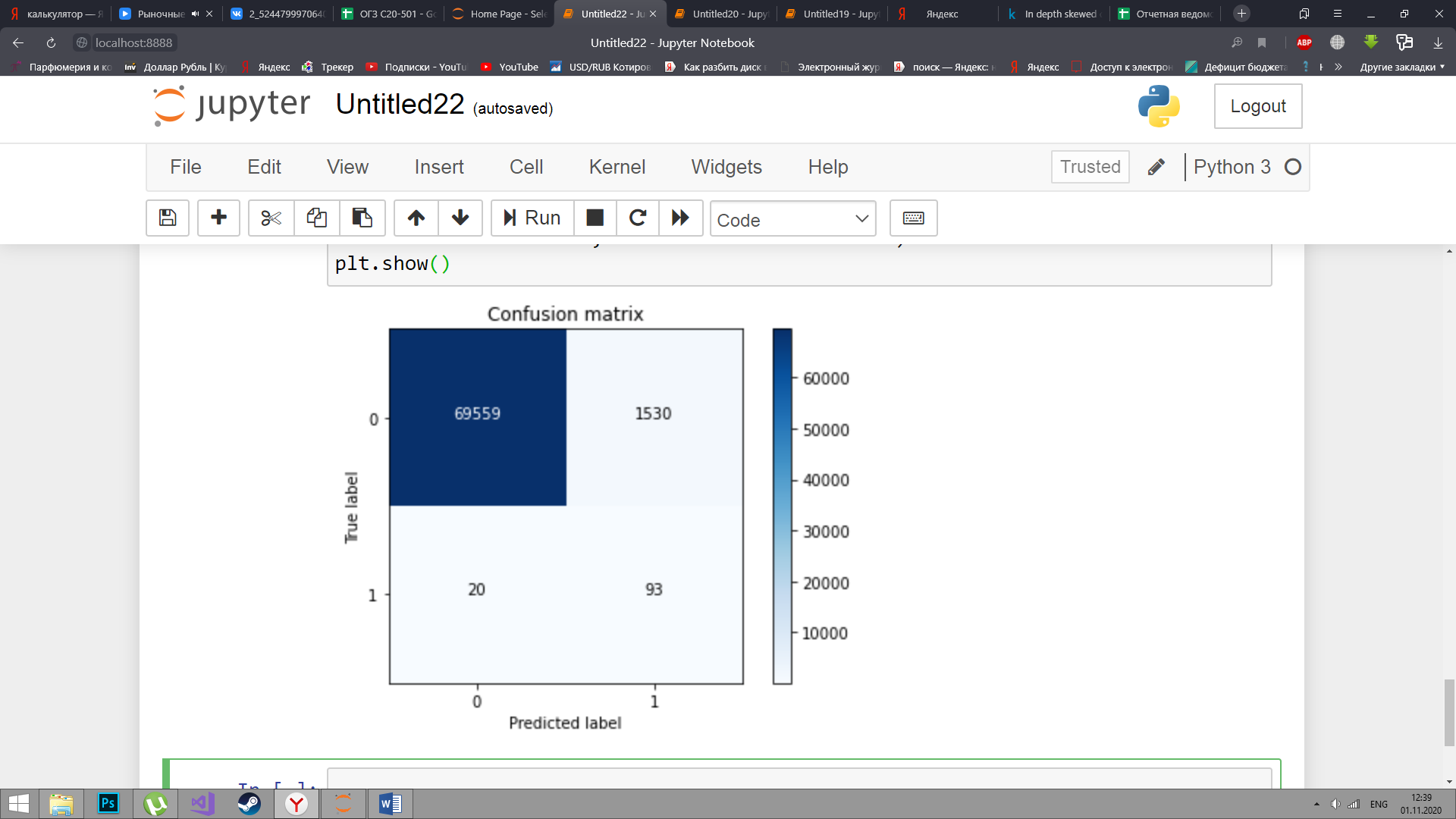
7 слой(decoder) – 30 нейронов, функция активации elu

Часть кода с построением модели показана на Fig.7.



1. Построение модели автокодировщика

Байесовский классификатор:



1. Матрица ошибок байесовского классификатора

* Среднее время оценки транзакции: 5,79048 мкс
* Точность на тестовом наборе: 97.8%
* F-мера: 0.988981

1. Показатели моделей

|  |  |  |  |
| --- | --- | --- | --- |
| **Модель** | **Точность на тестовом наборе %** | **F-мера** | **Среднее время оценки транзакции мкс** |
| НС Прямого распространения | 99,9 | 1,000 | 9,68 |
| Автокодировщик | 98,3 | 0,991 | 14,40 |
| Байесовский классификатор | 97,8 | 0,989 | 5,79 |

Все характеристики моделей вычислялись исходя из данных их матриц ошибок Fig. 4, Fig 6, Fig 8 по формулам (1) - (3).

# Discussion And Conclusions

В ходе данной работы на основе несбалансированного набора данных с банковскими транзакциями проводилось сравнение трёх алгоритмов машинного обучения, для выявления наиболее эффективного из них. В сравнении участвовали нейронная сеть прямого распространения, автокодировщик и наивный байесовский классификатор.

Для всех трех моделей были подобраны оптимальные характеристики. Были рассмотрены разные оптимизаторы, активационные функции и разные подходы к обработке датасета, был проведен анализ на результатах для каждого из подходов и выбраны лучшие из них, которые дают максимальную точность в обнаружении мошенничества. На данный момент наибольшая точность в обнаружении мошеннических транзакций среди моделей составляет 99,9%, а наибольшая скорость обработки транзакции составляет 5,79 мкс, что является хорошим показателем на текущий момент в сфере выявления мошенничества методами машинного обучения.

Из данных TABLE I можно сделать вывод, что нейронная сеть прямого распространения в данном сравнении показала наибольшую точность, но не самое быструю скорость обработки транзакций. При этом Байесовский классификатор показал существенно большую скорость обработки, в 1,7 раз быстрее нейронной сети прямого распространения и в 2,5 раз быстрее автокодировщика. Понятно, что данное время зависит от производительности системы, обрабатывающей транзакции, но отношение времен работы моделей останется неизменным вне зависимости от характеристик вычислительной машины. Именно поэтому Байесовский классификатор имеет преимущественное положение перед остальными моделями, т.к. его использование увеличит скорость работы системы примерно в 2 раза, а её точность упадет всего лишь на 2%.

Также по результатам сравнения F-мер моделей нейронная сеть прямого распространения показала результат близкий к абсолютному максимуму, что говорит о том, что в задачах выявления мошенничества эта модель является наилучшей с точки зрения качества оценки транзакций. Но при этом F-мера байесовского классификатора, хоть и является худшей среди трёх моделей, но всё же позволяет говорить о том, что данная модель также обладает достаточной полнотой и точностью чтобы использоваться в реальных системах выявления мошенничества с банковскими картами.

К недостаткам байесовского классификатора можно отнести самый большой процент ложных срабатываний на легитимных транзакциях. И действительно, представьте, что вы являетесь держателем карты, и после того, как вы совершили перевод, ваша карта блокируется, потому что по алгоритму банка ваша покупка была мошеннической. Вот почему мы не должны делать акцент только на выявлении случаев мошенничества, но и на количестве ложных срабатываний.

Очевидно, что для крупных организаций, у которых есть возможность иметь неограниченные ресурсы в производительности, использование байесовского классификатора или других алгоритмов, не использующих глубокое обучение, но обладающих сопоставимой точностью, может быть также крайне эффективно. Ведь они всё равно получат выигрыш в скорости работы.

В качестве дальнейшего продолжения исследования может быть рассмотрен вариант разработки нейросети, которая сама будет сравнивать качество оценки транзакции другими нейросетями. При использовании данной нейросети можно провести эксперимент аналогичный нашему, только с гораздо большим количеством моделей, что позволит получить более точные результаты и сделать более объективные выводы. Для данного сравнения рекомендуется взять такие алгоритмы машинного обучения, как случайный лес, логистическая регрессия, машина опорных векторов и метод к-ближайших соседей. Все эти алгоритмы также используются в задачах классификации транзакций и способны показывать довольно высокую точность. И поэтому именно они могут показывать наилучшую производительность в задачах мониторинга банковского мошенничества.

Также в целях увеличения точности моделей могут быть использованы различные техники по передискретизации несбалансированного набора данных, например, такие как under-sampling и over-sampling, то есть с уменьшением мажоритарного класса и увеличением миноритарного класса соответственно.

##### References

[1] I. Novikova, R. Muller, and T. Vostrova, “Противодействие мошенничеству: какие меры принимают компании?,” *Pwc*, 2018.

[2] Банк России, “Обзор операций, совершенных без согласия клиентов финансовых организаций за 2019 год,” 2019.

[3] E. V. Ilinich, “Мошеннические операции с банковскими пластиковыми картами как угроза экономической безопасности в сфере банковской деятельности,” pp. 41–45, 2013.

[4] Altexsoft, “Fraud Detection: How Machine Learning Systems Help Reveal Scams in Fintech, Healthcare, and eCommerce,” pp. 1–24, 2017.

[5] Козодаева О.Н. and Обыденнова А.С., “Способы совершения мошенничесва с использованием банковских карт,” pp. 52–58, 2017.

[6] КОЗЛОВ Д.В. and ДРУЖИНИНА Л.В., “Современные методы защиты от мошенничесикх угроз при эксплуатации платежных приложений,” pp. 1–4, 2017.

[7] Оголихин И.Д., “Обнаружение мошенничества с банковскими картами методами машинного обучения,” pp. 1–72, 2019.

[8] И. В. Воронов, Е. А. Политов, and В. М. Ефременко, “Обзор Типов Искусственных Нейронных Сетей И Методов Их Обучения,” *Вестник Кузбасского Государственного Технического Университета*, no. 3, pp. 38–42, 2007.

[9] В. А. Филиппов, Б. А. Щукин, and Л. Е. Муханов, “Мониторинг банковских систем с использованием простого классификатора байеса,” Moscow, pp. 318–323, 2018.

[10] M. A. Al-Shabi, “Credit Card Fraud Detection Using Autoencoder Model in Unbalanced Datasets,” *J. Adv. Math. Comput. Sci.*, vol. 33, no. 5, pp. 1–16, 2019, doi: 10.9734/jamcs/2019/v33i530192.