

Санкт-Петербургский государственный университет

ТОКАРЕВА Ирина Олеговна

Выпускная квалификационная работа

***Применение искусственных нейронных сетей в задаче прогнозирования
опасных конвективных явлений***

Уровень образования: магистратура

Направление: 02.04.02 «Фундаментальная информатика и информационные
технологии»

Основная образовательная программа: ВМ.5502 «Вычислительные
технологии»

Научный руководитель:
кандидат физ.-мат. наук, доцент,
Станкова Е. Н.

Рецензент:
кандидат физ.-мат. наук, с.н.с.,
Довгалюк Ю.А.

Санкт-Петербург

2020

Содержание

Введение	3
Постановка задачи	4
Обзор литературы	5
1 Численная модель конвективного облака	7
1.1 Процесс образования конвективных облаков	7
1.2 Классификация численных моделей конвективных облаков .	9
1.3 Нестационарная полуторамерная модель конвективного облака	11
2 Формирование данных для исследования	13
2.1 Реализация алгоритма получения данных	13
2.2 Данные	15
2.3 Предварительная обработка данных	16
3 Применение нейронных сетей	18
3.1 Многослойный персептрон	18
3.2 Персептронный комплекс	21
3.3 Сеть радиально-базисных функций	25
3.4 Вероятностная нейронная сеть	29
Выводы	31
Заключение	34
Список литературы	35

Введение

В настоящее время актуальна проблема прогнозирования таких опасных конвективных явлений, как грозы, град и обильные осадки, в связи с масштабом разрушений, которые они производят. Информация о наступлении подобных явлений нужна в первую очередь аэропортам, авиакомпаниям и службам МЧС.

Для прогнозирования таких явлений повсеместно используются современные численные модели конвективных облаков. Существует множество моделей, отличающихся как степенью детализации описания микрофизических процессов, так и размерностью.

Для научных исследований наибольший интерес представляют трехмерные модели с подробным описанием микрофизических и электрических характеристик облака, которые с большой степенью детализации описывают процессы в облаках во всей сложности их взаимодействия и, следовательно, должны обеспечивать наилучшее качество прогнозов. Однако их использование для оперативного прогнозирования в небольших метеорологических центрах, например, в метеорологических центрах аэропортов, невозможно из-за отсутствия там необходимой вычислительной мощности, которая нужна для проведения расчетов по таким моделям. Использование моделей меньшей размерности и более низкой функциональности ставит проблему определения вероятности развития грозы только путем анализа рассчитанных значений динамических и микрофизических характеристик облака, поскольку такие модели не имеют блока, описывающего электрические процессы.

В настоящее время методы машинного обучения считаются одним из наиболее перспективных инструментов для установления связи между выходными данными таких численных моделей и вероятностью появления грозы, таким образом они являются эффективным инструментом для прогнозирования опасных конвективных явлений.

В данной выпускной квалификационной работе рассмотрена возможность применения для решения задачи прогнозирования опасных конвективных явлений одного из видов машинного обучения, нейронных сетей, на примере прогнозирования грозы.

Постановка задачи

Цель работы: разработка и реализация алгоритма обработки результатов численного моделирования конвективных облаков с использованием нейронных сетей.

Для достижения поставленной цели были сформулированы следующие задачи:

1. на основе анализа различных моделей конвективного облака, выбрать модель, позволяющую реализовать прогнозирование эволюции облака в оперативном режиме;
2. реализовать получение и обработку входных данных для модели;
3. реализовать предварительную обработку выходных данных модели, которые будут использованы для выполнения прогноза;
4. реализовать прогнозирование опасных конвективных явлений (грозы) с использованием нескольких типов нейронных сетей;
5. для каждого типа нейронной сети найти структуру, с помощью которой получаются наилучшие результаты;
6. на основе анализа результатов сделать вывод о том, какая нейронная сеть является наилучшим вариантом для прогнозирования грозы.

Обзор литературы

В процессе написания выпускной квалификационной работы были проанализированы различные источники, описывающие основные положения физики конвективного облака, численные модели конвективных облаков и различные типы нейронных сетей.

Учебно-методическое пособие А. В. Назаренко «Опасные природные явления. Часть III. Опасные явления погоды конвективного происхождения» [1] помогло мне понять процесс образования разных видов конвективных облаков, а также ознакомиться с различными методами прогнозирования таких опасных конвективных явлений, как ливень, град, гроза, шквал и смерч. В пособии описаны такие методы прогнозирования грозы, как методы Вайтинга, Лебедевой и Пескова, которые основываются на расчете различных коэффициентов с использованием некоторых параметров конвективного облака.

Первая глава диссертации Н. О. Раба «Программный комплекс компьютерного исследования атмосферных процессов для многоядерных процессоров» [2] помогла ознакомиться с существующими численными моделями конвективных облаков и понять, как работает используемая в данной работе нестационарная полуторамерная модель.

Статья P. Jeatrakul, K. W. Wong «Comparing the Performance of Different Neural Networks for Binary Classification Problems» [3] познакомила меня с различными типами нейронных сетей, используемых в задачах бинарной классификации. Здесь рассматриваются пять типов сетей и сравниваются полученные с их помощью точности в задаче бинарной классификации с использованием трех разных наборов данных. На основе полученных здесь результатов к нейронным сетям, рассматриваемым в данной работе, были добавлены радиально-базисная и вероятностная сети.

Вторая и третья главы учебного пособия О. П. Солдатовой «Нейроинформатика» [4] помогли расширить знания о многослойном персептроне и радиально-базисных нейронных сетях.

С помощью статьи С. В. Коробковой «Проблемы эффективной аппроксимации многомерных функций с помощью нейронных сетей» [5] я ознакомилась с различными методами, помогающими повысить эффективность нейронных сетей при использовании малого обучающего набора. Из

этой статьи в данной работе используется метод перекрестной вариации. Еще один метод, помогающий при работе с ограниченным объемом данных, состоит в использовании перцептронного комплекса. Он описан в статьях авторов С. П. Дударова, А. Н. Диева, Н. А. Федосовой, Э. А. Кольцовой «Нейросетевое моделирование на основе перцептронных комплексов при малых объемах обучающих выборок» [6] и «Моделирование свойств конструкционного композитного материала, армированного углеродными нанотрубками, с использованием перцептронных комплексов» [7].

Статья Р. Н. Кветного, В. В. Кабачего, О. О. Чумаченко «Вероятностные нейронные сети в задачах идентификации временных рядов» [8] помогла лучше разобраться в том, как работает вероятностная нейронная сеть.

1 Численная модель конвективного облака

Облака являются одним из наиболее значимых факторов формирования погоды. Они являются существенной частью водного круговорота Земли, играют большую роль в радиационном режиме нашей планеты и служат источником таких опасных атмосферных явлений, как грозы и град, а также шквалы и смерчи, которые нередко приводят к человеческим жертвам и приносят значительный ущерб из-за своей разрушительной силы.

1.1 Процесс образования конвективных облаков

Облака – это взвешенные в атмосфере продукты конденсации водяного пара. Существует несколько типов облаков, отличающихся по внешнему виду и внутреннему строению:

- конвективные облака,
- волнистые облака,
- облака восходящего скольжения,
- облака турбулентного перемешивания.

Основная причина образования конвективных облаков – это конвекция, которая является одним из видов вертикальных движений в атмосфере. Если на частицы действуют силы плавучести, возникает термическая (свободная) конвекция, в случае воздействия механических сил возникающая конвекция называется вынужденной (динамической) [1].

Вынужденная конвекция в основном появляется на наветренных склонах возвышенностей и гор. В этом случае до достижения уровня конденсации воздух поднимается за счет натекания потока на возвышенность или упорядоченных вертикальных движений, а выше уровня конденсации – за счет выделения теплоты конденсации.

Термическая конвекция возникает в результате неустойчивой температуры в нижних слоях стратификации и неравномерного прогрева подстилающей поверхности. В результате появляются перегретые массы воздуха, которые начинают перемещаться по вертикали вверх, причем с высотой температура воздуха падает медленнее, чем температура окружающей среды. Таким образом он оказывается теплее и легче относительно окружающего воздуха и продолжает движение вверх. Если воздух на исходном

уровне является ненасыщенным, его температура в результате подъема меняется по сухоадиабатическому закону.

Из-за понижения температуры происходит конденсация водяного пара и воздух становится насыщенным. Далее подъем воздуха сопровождается выделением теплоты конденсации и сублимации, вследствие чего его температура понижается медленнее, чем при подъеме по сухой адиабате. Условно считается, что выше уровня конденсации воздух поднимается по влажной адиабате. Процесс подъема завершается тогда, когда температура поднимающегося и окружающего воздуха сравниваются.

По степени развития различают четыре категории конвективных облаков:

- плоские кучевые,
- средние кучевые,
- мощные кучевые,
- кучево-дождевые.

Наиболее опасные явления погоды связаны с кучево-дождевыми облаками. Например, такие облака могут быть очень опасны для пролетающих рядом самолетов. Восходящие и нисходящие движения в них могут вызывать сильные перегрузки и потерю управления. В грозовых кучево-дождевых облаках существует возможность поражения самолетов молниями. В кучево-дождевых облаках и на расстоянии нескольких километров от них наблюдается сильная турбулентность. С кучево-дождевыми облаками также связано появление большинства опасных атмосферных явлений.

Период жизни кучево-дождевых облаков включает три стадии [9]:

- формирование облака,
- максимальная фаза развития облака,
- разрушение облака.

Формирование кучево-дождевого облака начинается с момента возникновения облака и продолжается до выпадения осадков. Далее наступает максимальная стадия развития облака, в которую наблюдается появление в верхней части облака ледяных кристаллов и выпадение ливневых осадков. В облаках могут находиться снежинки, градины и крупные дождевые капли. Электрическое поле, формирующееся внутри облака, имеет большую напряженность, в результате чего возникают молнии. Стадия разрушения

является завершающей стадией в жизни облака.

1.2 Классификация численных моделей конвективных облаков

Численное моделирование является одним из наиболее эффективных инструментов для изучения конвективных облаков и исследования происходящих в них процессов. Существует множество моделей, отличающихся как степенью детализации описания микрофизических процессов, так и размерностью.

В зависимости от размерности, численные модели можно разделить на:

- одномерные,
- полуторамерные,
- двумерные,
- трехмерные.

Схематичное представление моделей разных размерностей представлено на рисунке 1.

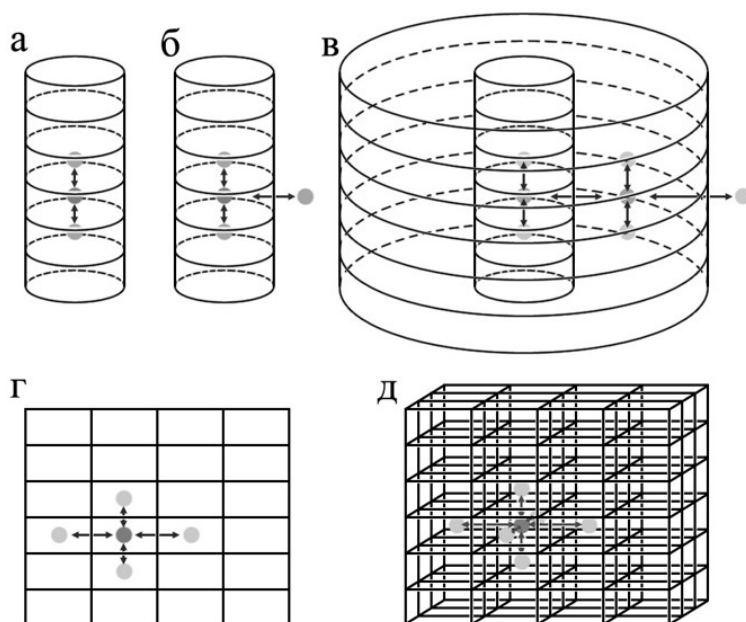


Рис. 1: Классификация моделей по размерности пространства (а – одномерная модель, б – полуторамерная модель с одним цилиндром, в – полуторамерная модель с двумя цилиндрами, г – двумерная модель, д – трехмерная модель).

Одномерные модели представляют облако в виде цилиндра, который разбивается на слои по высоте. В каждом слое происходит усреднение зна-

чений давления, температуры и других параметров облака. В данном представлении скорость потока имеет только вертикальную составляющую, и все процессы взаимодействия происходят только в вертикальном направлении.

В полутаромерных моделях облако также представляется в форме цилиндра, но в них дополнительно учитывается внешняя среда и присутствует радиальная составляющая скорости в горизонтальном направлении. Таким образом в этих моделях кроме взаимодействия в вертикальном направлении происходит взаимодействие между цилиндром и окружающей средой в горизонтальном направлении. Однако в таких моделях отсутствует нисходящий поток, который должен компенсировать развитие восходящего потока в неподвижной окружающей среде. При некоторых экспериментальных данных такое представление не способно адекватно воспроизвести все стадии жизни облака, в частности, стадию его разрушения, так как после прохождения облаком стадии развития ни динамические, ни микрофизические характеристики не претерпевают существенных изменений, и облако фактически стабилизируется.

Для решения этой проблемы была создана полутаромерная модель с двумя цилиндрами, в которой восходящий поток компенсируется путем добавления внешнего цилиндра. Внутренний цилиндр в данном представлении соответствует области с восходящим движением, где и происходит развитие облака, а внешний представляет безоблачную область с компенсирующим нисходящим движением.

В двумерных моделях вертикальное сечение пространства делится на прямоугольные области. Взаимодействие в них происходит по направлениям внутри этого сечения.

В трехмерных моделях пространство разбивается на области в форме прямоугольного параллелепипеда. Взаимодействие осуществляется по всем направлениям.

По способу описания микрофизических процессов модели конвективных облаков разделяют на две группы:

- модели с параметризованной микрофизикой,
- модели с детальной микрофизикой,

В моделях с параметризованной микрофизикой не учитываются спек-

тры частиц. Процессы испарения, конденсации, нуклеации и т.д., которые описывают переходы между паром, облачными каплями и дождевыми каплями, здесь задаются параметрически и не учитывают реальные размеры капель, а только их концентрацию. В таких моделях эволюция облака рассчитывается с использованием небольшого количества уравнений, поэтому эти модели являются вычислительно эффективными.

В моделях с детальной микрофизикой обычно учитываются водяные капли и ледяные частицы нескольких типов. Микрофизические процессы описываются более подробно и учитывают размеры, а иногда и форму частиц. Такие модели наиболее полно отражают процессы, происходящие в облаке, и соответственно позволяют делать более точные прогнозы, чем модели с параметризованной микрофизикой.

1.3 Нестационарная полутримерная модель конвективного облака

Для научных исследований наибольший интерес представляют трехмерные модели с детальной микрофизикой, которые наиболее полно описывают природные процессы в облаках и, следовательно, должны обеспечивать наилучшее качество прогнозов. Однако их использование для оперативного прогнозирования в небольших метеорологических центрах невозможно из-за отсутствия там необходимой вычислительной мощности. Поэтому для оперативного прогноза используются более простые модели, которые требуют меньше вычислительных ресурсов. Наилучшими моделями для этих целей являются полутримерные модели, способные предоставлять достоверные прогнозы на краткосрочный период [2].

В данной работе используется разработанная Н.О. Раба и Е.Н. Станковой полутримерная нестационарная модель с двумя цилиндрами и с подробным описанием микрофизических процессов для жидкой и твердой фаз [2]. Схематичное изображение модели показано на рисунке 2. В данной модели рассматриваются капельки воды, а также ледяные частицы различных типов: снежинки, крупа, градины и 3 типа кристаллов.

Модель состоит из динамического и микрофизического блоков. Динамический блок модели представляет собой систему дифференциальных

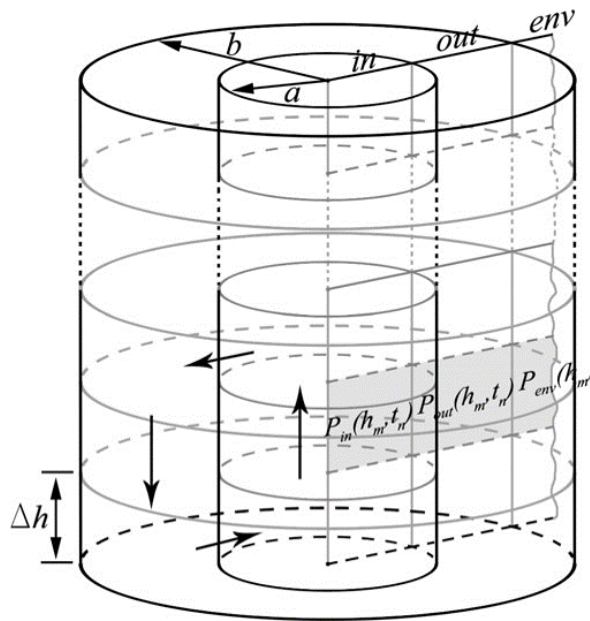


Рис. 2: Полутримерная модель конвективного облака, разработанная Н.О. Раба и Е.Н. Станковой [2].

уравнений в частных производных, которые описывают эволюцию во времени и пространстве значений таких параметров конвективного облака, как отношение смеси водяного пара, капель воды и ледяных кристаллов, температура и вертикальная скорость.

Микрофизический блок модели состоит из системы кинетических уравнений, описывающих изменение спектров капель и кристаллов под действием процессов конденсации на поверхности частиц, испарения с поверхности частиц, коагуляции, нуклеации, замерзания капелек и таяния ледяных частиц. Для этого в каждом слое для каждого цилиндра хранится информация о распределениях по размерам капель, ледяных частиц каждого типа и частиц аэрозоля. Капли разных размеров имеют разные устоявшиеся скорости падения относительно потока. Чем больше размер капли, тем больше скорость падения. Маленькие капли, скорость падения которых почти равна нулю, образуют облако, крупные капли выпадают в виде осадков. Предполагается, что скорость падения частиц аэрозоля относительно потока равна нулю.

2 Формирование данных для исследования

Исследования с помощью методов машинного обучения, к которым относятся нейронные сети, основываются на данных, поэтому для получения наилучших результатов необходимо использовать достоверные источники информации для получения данных, сформировать их корректную структуру и провести предварительную обработку.

Изначально для исследований был предоставлен набор данных, полученных Д.А. Петровым с помощью разработанной им информационной системы интеграции метеорологической информации [10]. Однако этот набор имеет небольшой объем, а для проведения исследований с использованием нейронных сетей требуется значительное количество данных. Поэтому частью этой работы являлась реализация алгоритма получения дополнительных данных для исследования.

2.1 Реализация алгоритма получения данных

Первым шагом являлось получение информации о том, где и когда происходили различные метеорологические явления, с сайта Всероссийского научно-исследовательского института гидрометеорологической информации (<http://meteo.ru>). На сайте можно получить массив с данными об атмосферных явлениях, наблюдения за которыми входят в программу стандартных наблюдений на метеорологических станциях. На станции определяется вид атмосферного явления, время его начала и окончания, продолжительность и интенсивность явления. На сайте можно получить данные наблюдений с 1966 года.

Затем была написана программа на языке программирования Python, которая извлекает из полученной информации те данные, которые соответствуют наступлению грозы или отсутствию метеорологических явлений, с их помощью получает данные радиозондирования с сайта Университета Вайоминга (<http://weather.uwyo.edu>) и приводит полученные радиозондировки к формату входных данных численной модели.

Далее производится моделирование эволюции конвективного облака с использованием полуторамерной нестационарной численной модели. На рисунке 3 показан вид программы, которая производит расчет по модели.

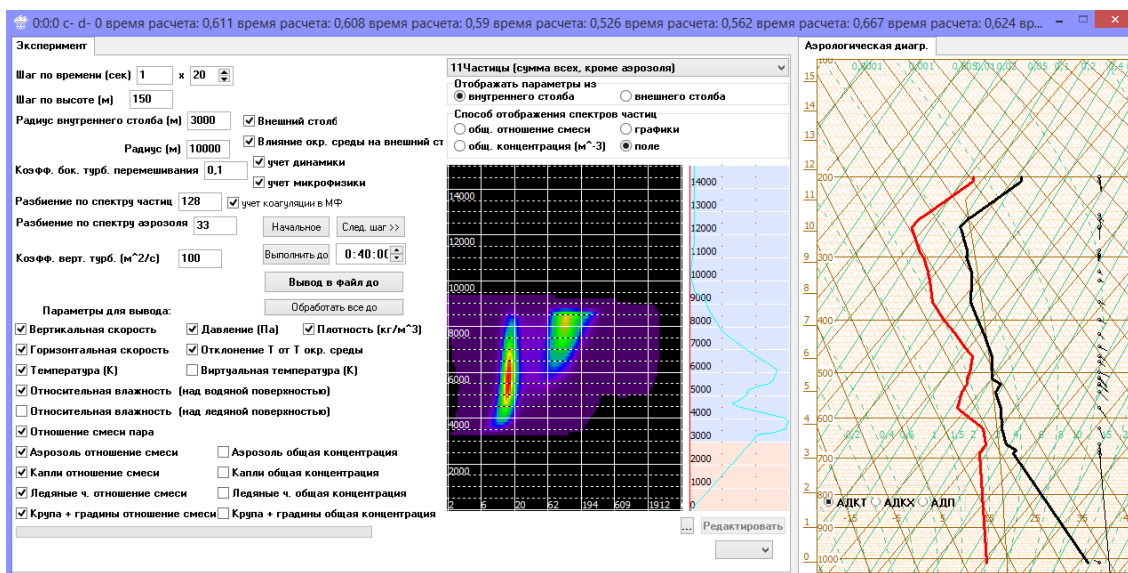


Рис. 3: Программа, которая осуществляет расчет по модели.

Справа представлена аэрологическая диаграмма, черная линия на которой показывает температуру воздуха, а красная – температуру точки росы, характеризующую влажность в атмосфере. В центре отображается поле развития облака за заданный промежуток времени. Слева представлены различные параметры модели, которые можно настроить перед ее запуском, в том числе и выводимые в файл динамические и микрофизические характеристики облака.

В данном случае выводятся следующие характеристики облака:

- вертикальная составляющая скорости,
- горизонтальная составляющая скорости,
- температура,
- отклонение температуры от температуры окружающей среды,
- относительная влажность (над водяной поверхностью),
- отношение смеси пара,
- давление,
- плотность,
- общее отношение смеси аэрозолей, капель, ледяных частиц, крупы и градин.

На рисунке 4 показан фрагмент выходного файла. В нем сохраняются значения характеристик для всего времени эволюции облака и различных значений высоты.

Последним шагом является расчет значений вертикальной мощности

Время(с),Высота(м),Название параметра,Значение параметра					
0.0,0.0,velocity,0					
0.0,0.0,velocityU,0					
0.0,0.0,temperature,294.21					
0.0,0.0,deltaTemperature,0					
0.0,0.0,relativeHumidity,1					
0.0,0.0,vapor,0.015348391					
0.0,0.0,pressure,101186.44					
0.0,0.0,density,1.1981394					
0.0,0.0,aerosol,6.0552032E-0007					
0.0,0.0,drop,0					
0.0,0.0,ice,0					
0.0,0.0,hailAndGrits,0					

Рис. 4: Фрагмент выходного файла численной модели.

облака и определение высоты и времени, соответствующих максимальному развитию и максимальной влажности облака. Значения параметров облака, соответствующие полученным высоте и времени, сохраняются в CSV файл и будут в последствии использоваться для прогнозирования грозы.

2.2 Данные

Сформированный набор данных содержит 416 примеров, из которых 220 соответствуют наличию опасного конвективного явления (грозы) и 196 – его отсутствию. Для каждого примера из набора был создан вектор меток, по которому можно однозначно определить то, к какому классу относится данный пример: (1, 0) для примеров, соответствующих наличию опасного конвективного явления, (0, 1) – для тех, которые соответствуют его отсутствию.

Для обработки данных с использованием нейронных сетей, их нужно разделить на три набора:

- тренировочный набор, с использованием которого происходит обучение нейронной сети;
- валидационный набор, который используется для проверки результатов тренировки сети во время обучения;
- тестовый набор, который используется для получения оценки сети после того, как модель полностью обучена.

Так как данные в нашем случае имеют небольшой объем, было решено использовать тестовую выборку в качестве валидационной. Соответ-

ственно данные были разделены на тестовую и обучающую выборки. Обучающая выборка содержит 333 примера, а тестовая – 83.

2.3 Предварительная обработка данных

Нейронные сети, как и все алгоритмы машинного обучения, существенно зависят от качества исходных данных. Поэтому прежде чем приступить к построению нейронной сети, нужно провести предобработку данных. Для этого была написана программа на языке программирования Python.

Сначала данные были нормализованы с помощью метода Standard Scaler из библиотеки scikit-learn, который преобразует их к стандартному нормальному распределению.

Затем были выбраны наиболее значимые признаки. Для этого использовался метод Recursive Feature Elimination из библиотеки scikit-learn с алгоритмом Random Forest в качестве классификатора. Метод заключается в следующем. Классификатор сначала обучается на полном наборе признаков, затем наименее важный признак исключается, и процедура рекурсивно повторяется все с меньшим и меньшим набором. На рисунке 5 показан получившийся после применения метода график точности классификации в зависимости от числа используемых признаков.

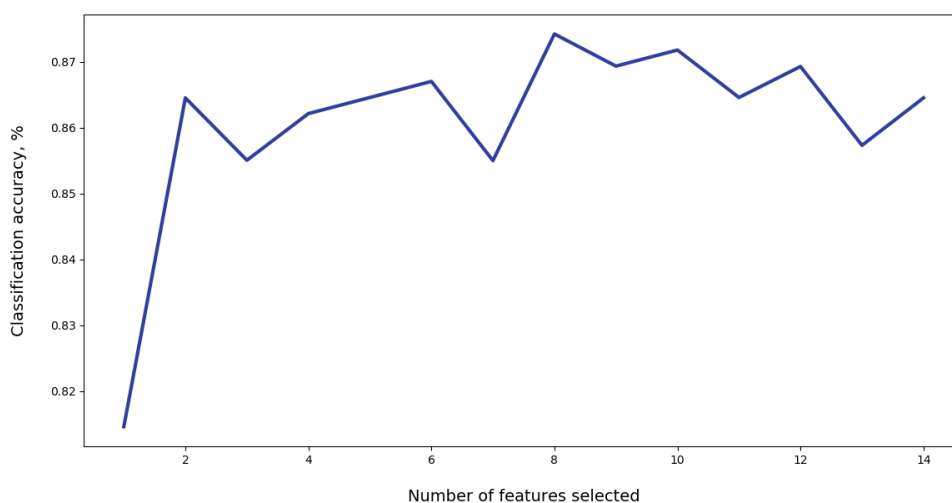


Рис. 5: График зависимости точности классификации от числа использованных признаков.

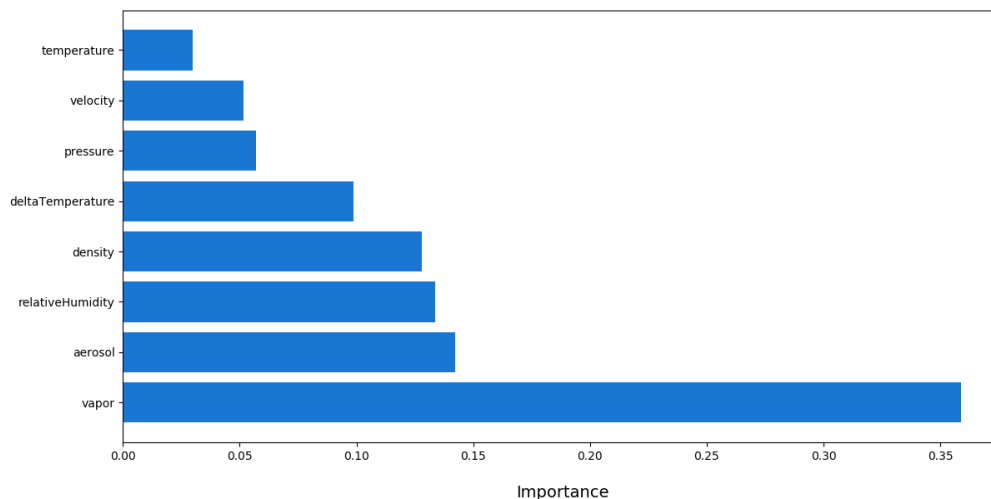


Рис. 6: Названия выбранных признаков и их важность в задаче классификации.

Как видно из рисунка, максимальная точность достигается при использовании восьми признаков. Этими признаками являются:

- вертикальная составляющая скорости,
- давление,
- плотность,
- температура,
- отклонение температуры от температуры окружающей среды,
- относительная влажность (над водяной поверхностью),
- отношение смеси пара,
- общее отношение смеси аэрозолей.

Их важность в задаче классификации показана на рисунке 6.

3 Применение нейронных сетей

Для прогнозирования опасных конвективных явлений в данной работе были использованы четыре типа нейронных сетей:

- многослойный персептрон,
- персептронный комплекс,
- сеть радиально-базисных функций,
- вероятностная нейронная сеть.

3.1 Многослойный персептрон

Многослойный персептрон является классической структурой, используемой в задачах классификации. Общая структура персептрона показана на рисунке 7.

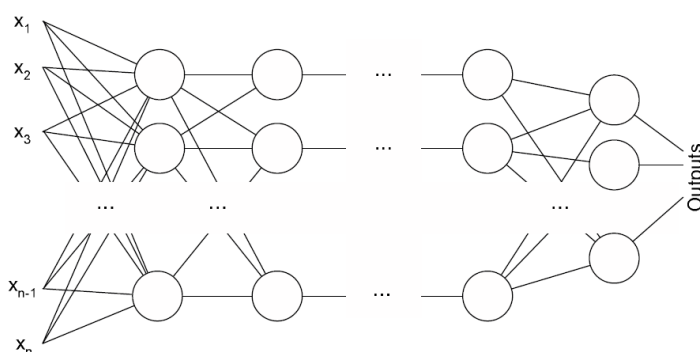


Рис. 7: Общая структура многослойного персептрона.

В статье [6] говорится, что соотношение объема обучающего набора и количества настраиваемых параметров сети (весов) является одним из факторов, влияющих на моделирующую способность персептрона. Если это соотношение близко к единице, персептрон будет способен только восстанавливать данные, не обладая при этом свойством обобщения. Если соотношение слишком большое, он будет усреднять данные без учета деталей. В связи с этим в большинстве случаев рекомендуется выбрать такую конфигурацию сети, чтобы это соотношение находилось в диапазоне от двух до пяти.

В книге [4] приводится теорема об универсальной аппроксимации, из которой следует, что для решения задач с помощью многослойного персептрона достаточно использования структуры с одним скрытым слоем.

Однако из теоремы не следует, что один скрытый слой является оптимальным в смысле качества обобщения данных или времени обучения. В практических задачах чаще всего используются перцептроны с одним или двумя скрытыми слоями.

Для определения оптимального количества скрытых слоев и числа нейронов в них была проведена серия экспериментов и построены графики зависимости точности и ошибки классификации от числа нейронов в одном скрытом слое (рисунок 8). Черная линия соответствует наличию одного скрытого слоя в сети, красная – двум. На графиках показаны усредненные значения точности и ошибки, полученных за 5 экспериментов.

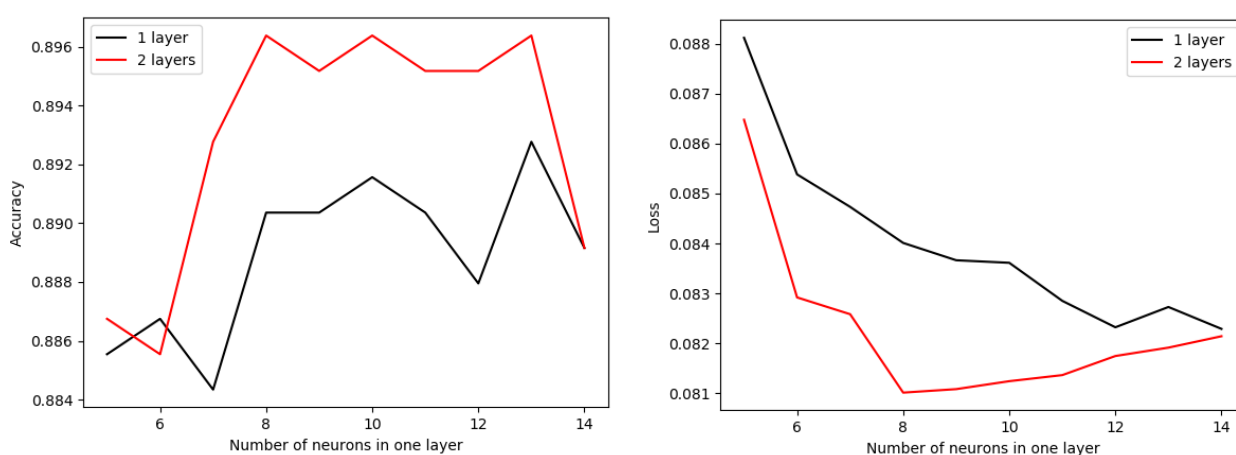


Рис. 8: Зависимость точности (слева) и ошибки (справа) классификации от числа нейронов в одном скрытом слое для многослойного перцептрона с одним и двумя скрытыми слоями.

Из графиков видно, что структура сети с двумя скрытыми слоями показывает лучшие результаты, причем наименьшее значение ошибки достигается при наличии восьми нейронов в каждом скрытом слое. Получившаяся структура сети показана на рисунке 9. В данной работе эта нейронная сеть реализована на языке программирования Python с использованием библиотеки Keras. После каждого скрытого слоя добавлены слой BatchNormalization и Dropout, функцией активации нейронов в скрытых слоях выбран гиперболический тангенс. В качестве функции потерь здесь используется среднеквадратическая ошибка. Точность классификации составила 89,1%, ошибка равна 0,082. Изменение точности и ошибки во время обучения показано на рисунке 10.

Соотношение объема обучающего набора и количества настраиваемых весов в сети составляет

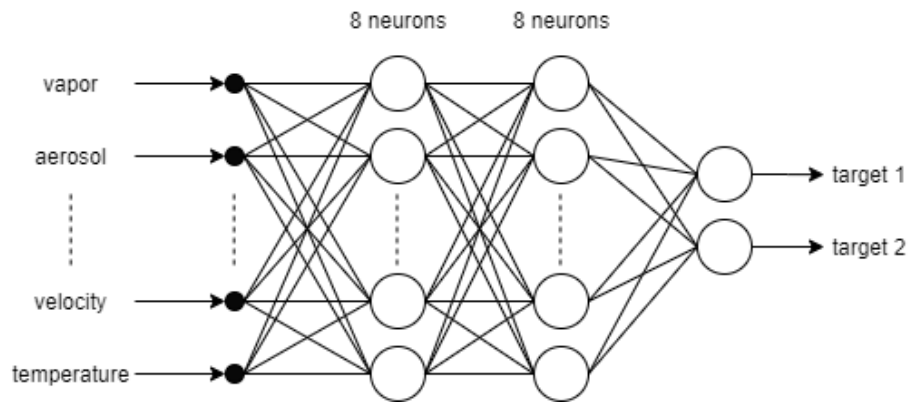


Рис. 9: Структура многослойного перцептрона, с помощью которой получена точность 89,1%..

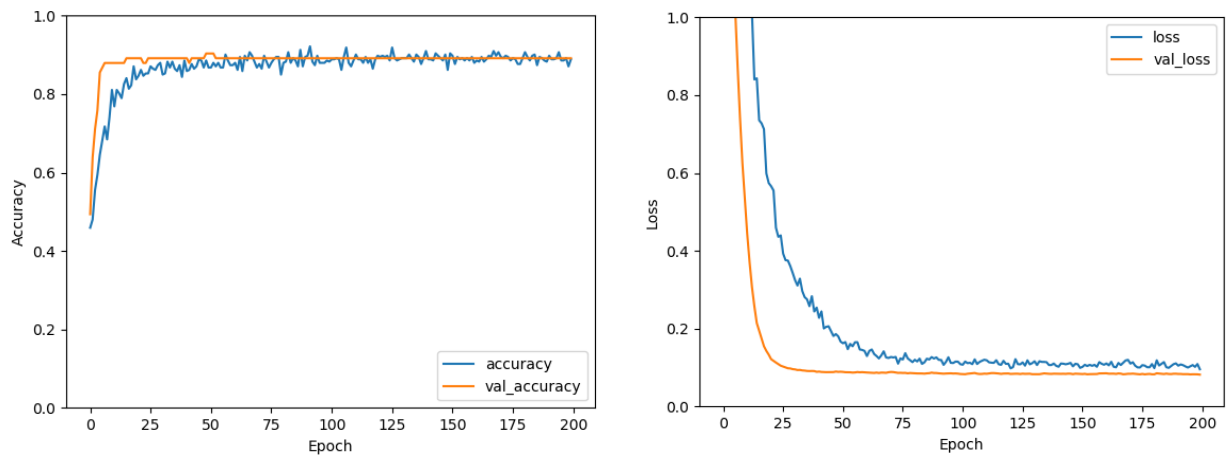


Рис. 10: Изменение точности (слева) и ошибки (справа) во время обучения для многослойного перцептрона.

$$\frac{333}{194} = 1,7,$$

что не входит в рекомендуемый интервал, однако находится достаточно близко к нему.

В статье [5] предлагается метод для более точной оценки тестирования сети при малом объеме обучающего набора, который называется методом перекрёстной вариации. Его суть состоит в том, что из выборки поочерёдно исключается один пример, который выступает в качестве тестового, и нейронная сеть обучается на оставшихся примерах. После обучения рассчитывается значение ошибки для тестового примера, а также выясняется, правильно ли был определен его класс. Затем исключенный пример возвращается в выборку, и процесс повторяется. Таким образом каждый пример в выборке выступает как в роли обучающего, так и в роли

тестового. После проведения расчетов для всех тестовых примеров, можно усреднить результаты и получить более точные оценки ошибки и точности сети. Недостатками данного метода являются необходимость многократного повторения обучающей процедуры, что занимает значительное количество времени, а также возможная неточность оценок отдельных ошибок тестовых примеров из-за влияния стохастической составляющей процесса обучения.

После применения этого метода ошибка сети получилась равной 0,080, а точность составила 89,9%.

3.2 Персептронный комплекс

Использование алгоритмов машинного обучения, основанных на нейронных сетях, может быть неэффективным при небольшом объеме обучающей выборки [6]. Поэтому, чтобы повысить эффективность построенной нейронной сети, был использован один из методов, который позволяет решить проблему ограниченного объема данных.

Этот метод описан в статье [7]. Суть метода состоит в разделении набора входных и выходных переменных сети на несколько персептронов с более простой структурой и последующим их объединении в персептронный комплекс. На рисунке 11 показана общая структура такого комплекса.

Алгоритм обучения персептронного комплекса [7]:

1. Для каждого персептрона первого уровня:
 - (а) Учитывая входные и выходные переменные текущего персептрона, на основе исходных данных формируются обучающая и тестовая выборки;
 - (b) Производится обучение персептрона;
 - (c) Вычисляются и сохраняются значения выходов для всех примеров обучающей и тестовой выборок.
2. Для результирующего персептрона:
 - (а) Учитывая входные и выходные переменные персептрона, на основе исходных данных и рассчитанных значений выходов персептронов первого уровня формируются обучающая и тестовая выборки;
 - (b) Производится обучение персептрона.

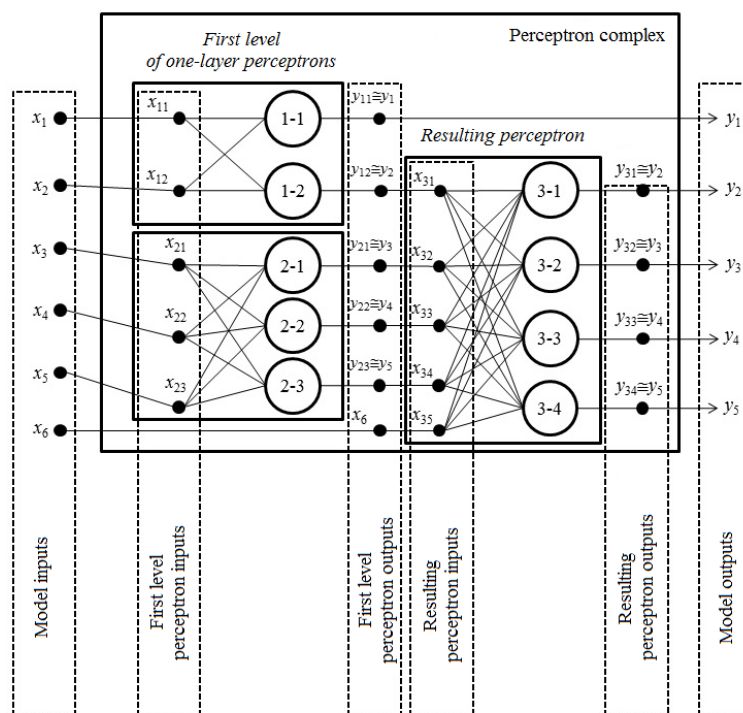


Рис. 11: Общая структура персептронного комплекса [7].

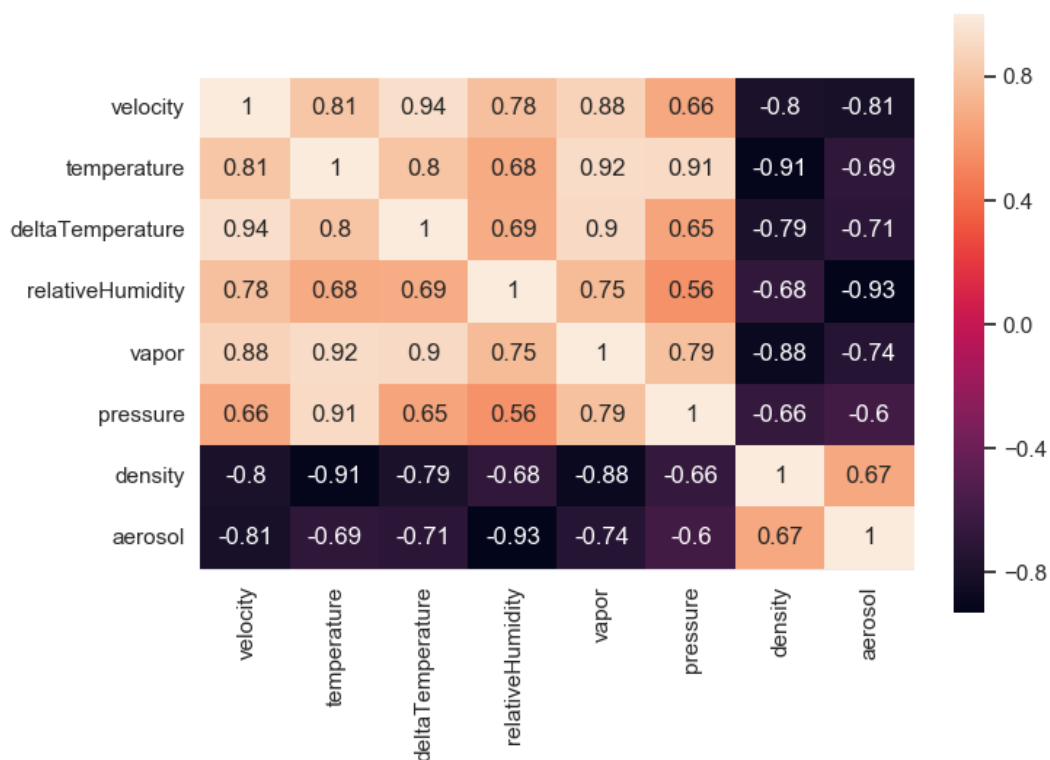


Рис. 12: Карта корреляции признаков.

Так как многослойный персептрон из предыдущего пункта (рисунок 9) имеет два выхода, персептронный комплекс будет содержать два персеп-

трона первого уровня, у каждого из которых будет по одному выходу. Все персептроны в комплексе являются персептронами с одним скрытым слоем, в которых после скрытого слоя присутствуют слой BatchNormalization и Dropout, функцией активации нейронов скрытого слоя выбран гиперболический тангенс. В качестве функции потерь здесь также используется среднеквадратическая ошибка. Персептронные комплексы были реализованы с использованием библиотеки Keras.

Сначала была рассмотрена структура персептронного комплекса без проскоков входа, в которой входами результирующего персептрона являются только выходы персептронов первого уровня. Набор признаков был разделен на две группы, которые использовались в качестве входных переменных персептронов первого уровня, таким образом, чтобы каждый из персептронов имел достаточно информации для независимого принятия решения о принадлежности исходного вектора к определенному классу. Для этого использовалась карта корреляции признаков, показанная на рисунке 12. Признаки, модуль корреляции между которыми превышает некоторое значение r , помещались в разные группы. Из карты видно, что за r можно взять значение $r = 0,88$. Также видно, что разделение на группы нельзя провести однозначно.

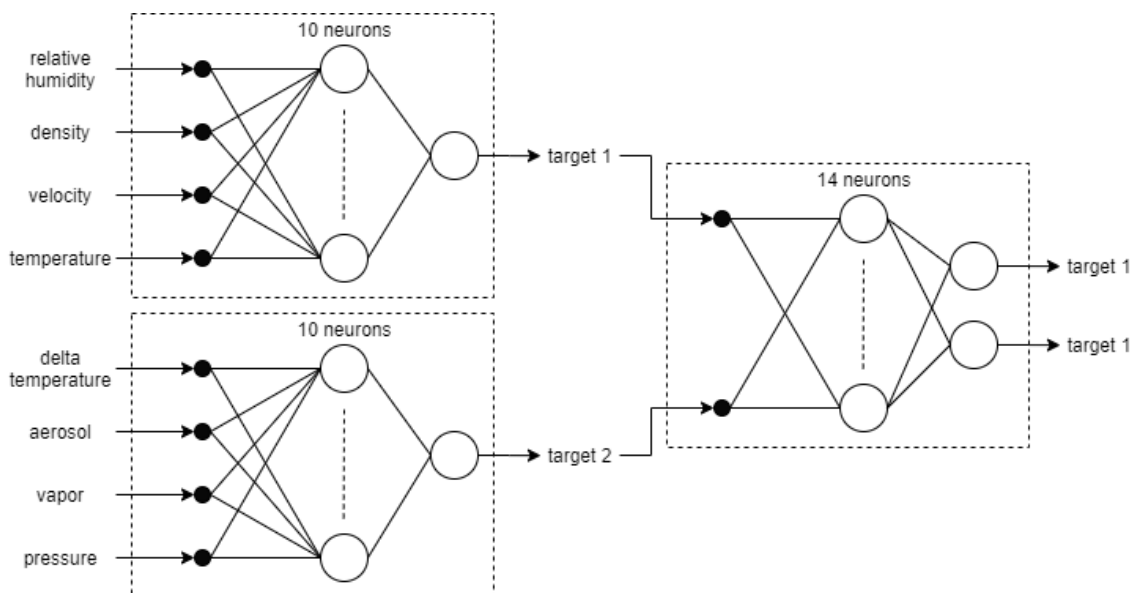


Рис. 13: Структура первого персептронного комплекса, с помощью которой получена точность 90,4%..

Были рассмотрены различные варианты разделения признаков на группы. Структура одного из комплексов, давших наилучший результат,

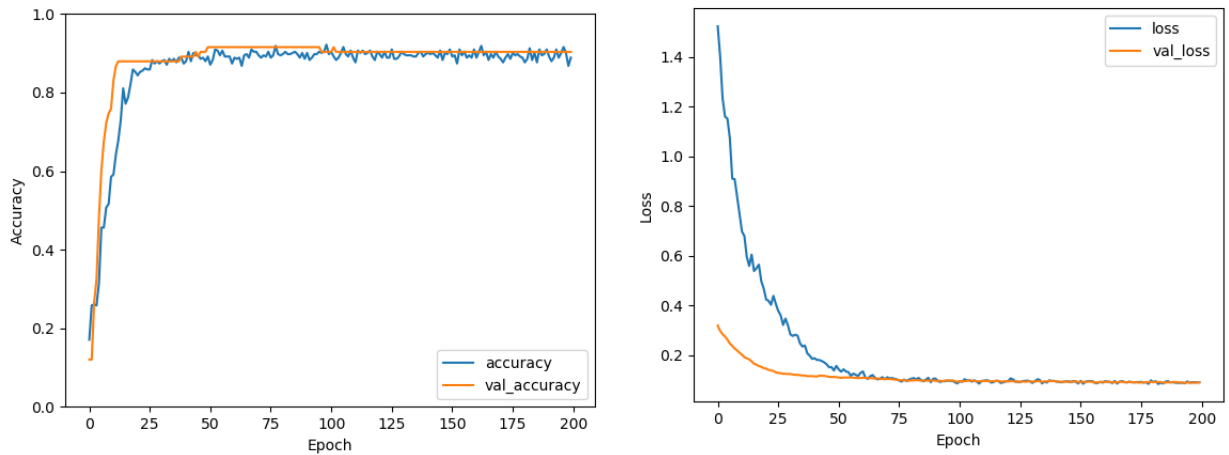


Рис. 14: Изменение точности (слева) и ошибки (справа) во время обучения для первого перцептронного комплекса.

показана на рисунке 13. Ошибка составила 0,081, точность – 90,4%. Изменение точности и ошибки во время обучения показано на рисунке 14.

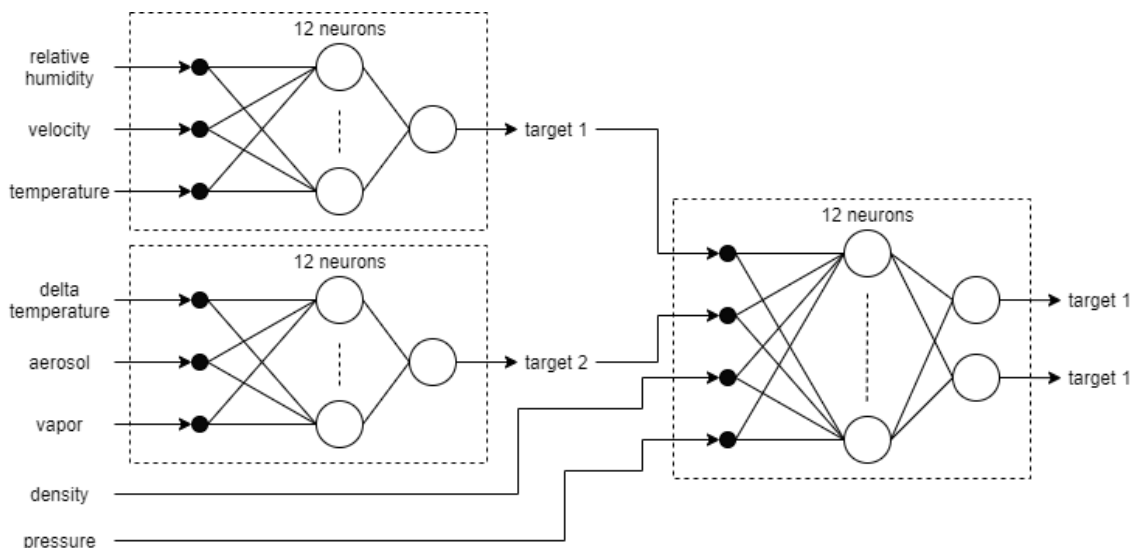


Рис. 15: Структура второго перцептронного комплекса, с помощью которой получена точность 91,6%..

В этом перцептронном комплексе результирующий перцептрон просто обрабатывает значения выходов, полученные с помощью перцептронов первого уровня. Для улучшения результата была рассмотрена структура комплекса с двумя проскоками входа, в которой входами результирующего перцептрона являются выходы перцептронов первого уровня и два признака из исходного набора. Таким образом результирующий перцептрон принимает решение о принадлежности исходного вектора к определенному классу, основываясь не только на информации, полученной от перцептронов первого уровня. Первым таким признаком была выбрана плотность,

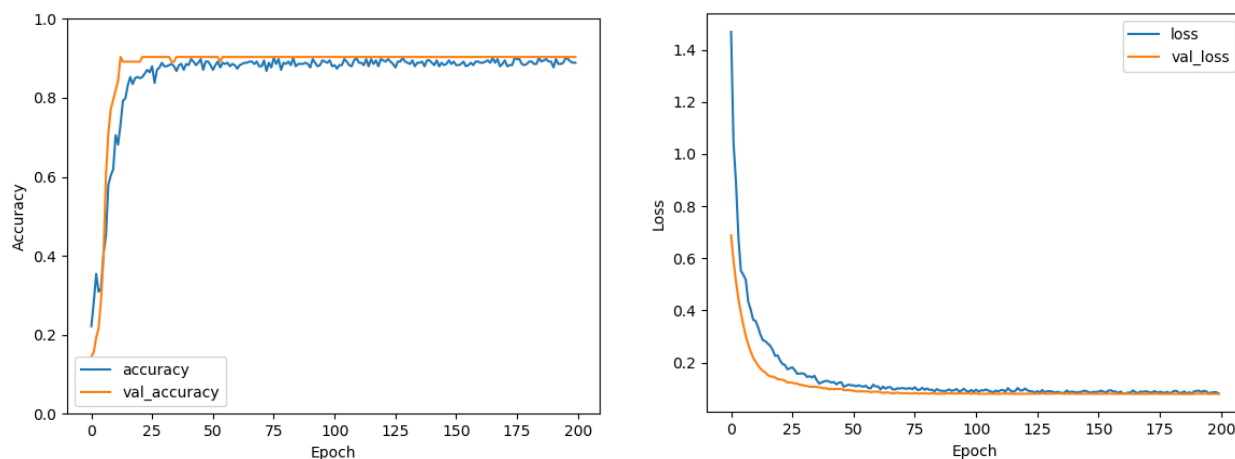


Рис. 16: Изменение точности (слева) и ошибки (справа) во время обучения для второго персептронного комплекса.

так как из карты коррелиции видно, что коррелиция между ней и температурой, с которой она находится в одной группе, составляет 0,91. Второй признак, давление, был выбран методом подбора. Получившаяся структура персептронного комплекса показана на рисунке 15. Получившаяся точность составила 91,6%, ошибка равна 0,078. Изменение точности и ошибки во время обучения показано на рисунке 16.

3.3 Сеть радиально-базисных функций

Сеть радиально-базисных функций (РБФ) имеет множество применений, включая аппроксимацию функций, прогнозирование временных рядов и классификацию.

Общая структура сети показана на рисунке 17. Сеть состоит из двух слоев: слоя нейронов РБФ (слоя образцов) и слоя суммирования, в котором содержится по одному нейрону для каждого класса.

Каждый нейрон РБФ хранит в себе вектор-прототип, который может являться одним из векторов из обучающего набора. Такие нейроны сравнивают входной вектор со своим прототипом и выдают значение от 0 до 1, которое является мерой сходства векторов. Существуют различные варианты задания функции подобия, но самый популярный основан на использовании распределения Гаусса, так как с помощью использования линейной комбинации гауссианов можно аппроксимировать любую функцию. Из этого вытекает преимущество такой сети перед, например, многослойным персептроном, состоящее в том, что она может смоделировать произволь-

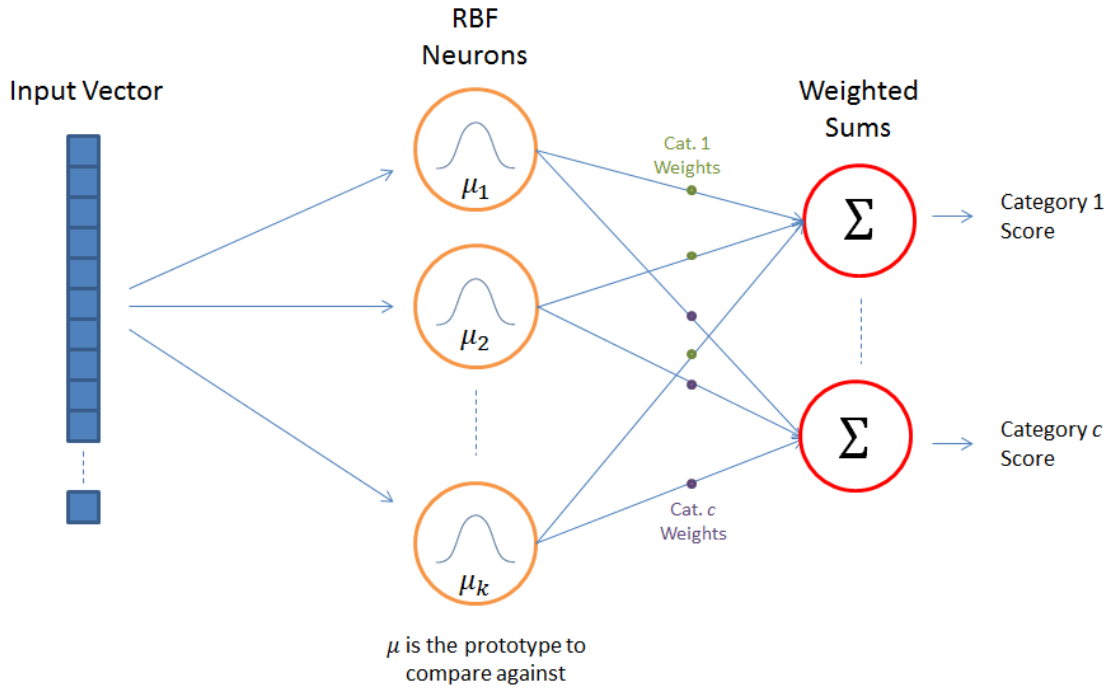


Рис. 17: Общая структура сети радиально-базисных функций [11].

ную нелинейную функцию с помощью всего лишь одного промежуточного слоя.

Ниже приведена формула распределения Гаусса в одномерном случае.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1)$$

где μ – среднее, σ – стандартное отклонение.

Функция активации нейрона РБФ выглядит немного по-другому [11]:

$$\varphi(x) = e^{-\beta\|x-\mu\|^2}, \quad (2)$$

где μ – это вектор-прототип. Эта функция называется радиально-базисной функцией. По сравнению с (1) здесь отсутствует коэффициент перед экспонентой, так как в сети присутствует выходной слой, на котором будет производиться взвешенное суммирование. Также здесь введен коэффициент β , который определяется как

$$\beta = \frac{1}{2\sigma^2}. \quad (3)$$

Таким образом в процессе обучения сети нам нужно определить три

набора параметров: векторы-прототипы и коэффициенты β для нейронов РБФ и веса для выходного слоя сети.

Векторы-прототипы могут быть определены различными способами. Например, можно случайным образом выбрать N примеров из тренировочного набора. Однако наиболее предпочтительным считается использование метода k средних. Суть метода состоит в разделении данных на заранее заданное число кластеров. Центры этих кластеров и будут использоваться в качестве векторов-прототипов. Причем перед применением метода k средних рекомендуется разделить тренировочный набор по классам, так как таким образом ни один кластер не будет содержать данные, относящиеся к разным классам.

Коэффициент β можно задать общим для всех РБФ нейронов или определить индивидуальные значения коэффициента для каждого нейрона. Если для задания векторов-прототипов используется метод k средних, то β можно определить по формуле (3), в которой стандартное отклонение будет равняться среднему расстоянию между всеми векторами в кластере и его центром:

$$\sigma = \frac{1}{m} \sum_{i=1}^m \|x_i - \mu\|,$$

где μ – это центр кластера, x_i – тренировочный пример, входящий в кластер, m – число примеров в кластере.

Последний набор параметров, веса, можно определить с помощью метода градиентного спуска. Однако в данной работе эта нейронная сеть реализована с использованием библиотеки Keras, поэтому для определения весов используется алгоритм обратного распространения ошибки. Слой РБФ нейронов реализован с использованием кода из GitHub репозитория [12], где можно сделать обучаемыми параметры β и векторы-прототипы. Таким образом в данной работе перед обучением сети задаются начальные значения векторов-прототипов и коэффициентов β , а затем они вместе с весами корректируются в процессе обучения. В качестве функции потерь здесь используется средняя квадратическая ошибка.

Для построения оптимальной структуры сети, а также для выбора лучшего способа задания векторов-прототипов были построены графиче-

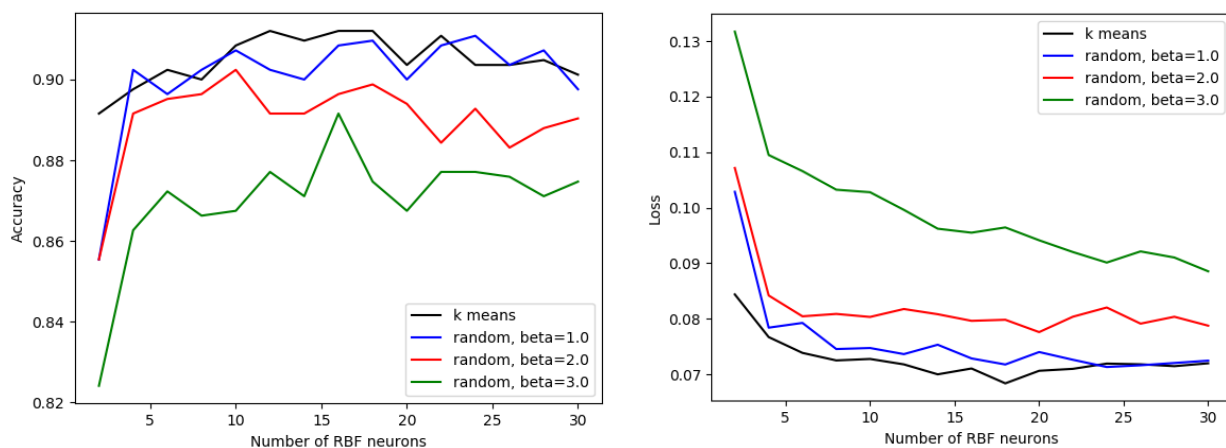


Рис. 18: Зависимость точности (справа) и ошибки (слева) классификации от числа РБФ нейронов при разных методах выбора векторов-прототипов и различных значениях параметра β .

ки зависимости точности и ошибки классификации от числа нейронов в слое РБФ (рисунок 18). Черная линия на графиках соответствует выбору векторов-прототипов с использованием метода k средних, коэффициенты β рассчитаны по формуле (3). Остальные три линии соответствуют случайному выбору векторов-прототипов и различным значениям коэффициента β , которые задавались общими для всех нейронов РБФ. На графиках показаны усредненные значения точности и ошибки, полученных за 10 экспериментов.

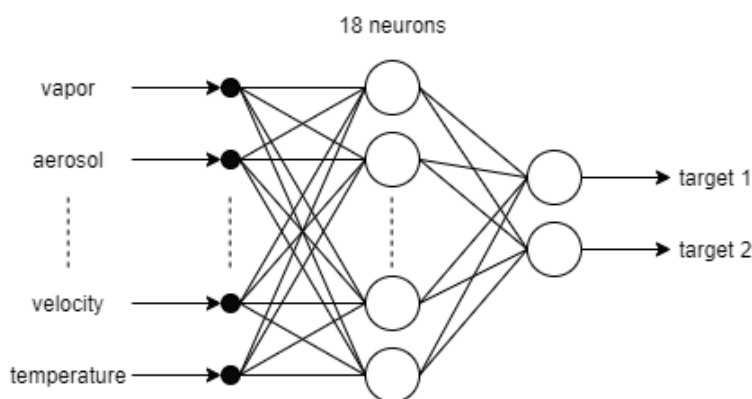


Рис. 19: Структура радиально-базисной нейронной сети, с помощью которой получена точность 91,6%.

Из графика точности видно, что наибольшую точность классификации дает применение метода k средних для нахождения начальных векторов-прототипов и наличие в сети от 12 до 18 РБФ нейронов. График ошибки показывает, что минимальная ошибка классификации достигается при числе РБФ нейронов, равном 18.

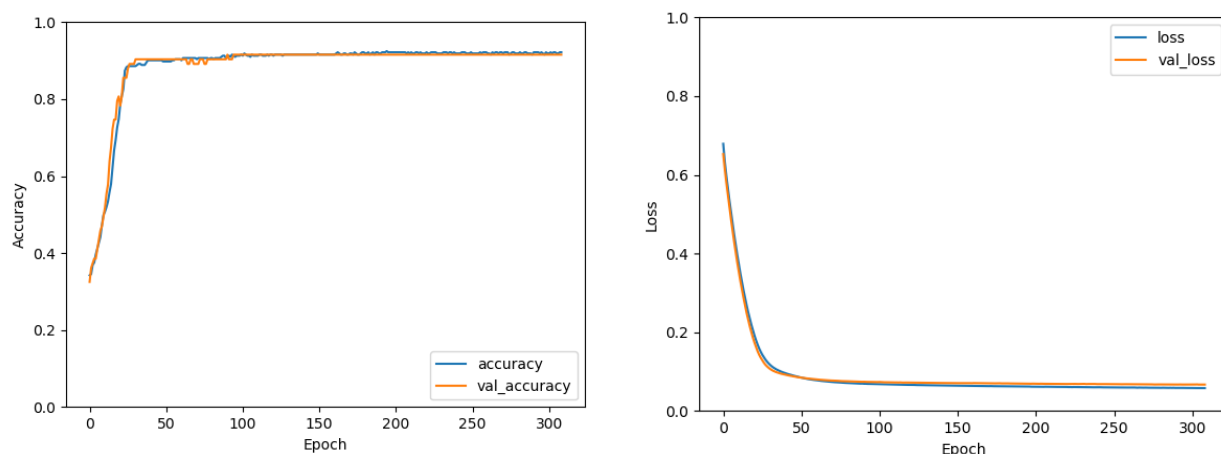


Рис. 20: Изменение точности (слева) и ошибки (справа) во время обучения для радиально-базисной нейронной сети.

Получившаяся структура сети показана на рисунке 19. Точность классификации составила 91,6%, ошибка равна 0,069. Изменение точности и ошибки во время обучения показано на рисунке 20.

3.4 Вероятностная нейронная сеть

Вероятностная нейронная сеть относится к сетям радиально-базисного типа, она широко используется в задачах распознавания образов и классификации. Структура этой сети (рисунок 21) немного отличается от структуры сетей РБФ [8].

1. Количество нейронов в слое образцов здесь равно числу примеров в обучающем наборе, то есть сеть запоминает весь обучающий набор;
2. Каждый нейрон слоя суммирования связан только с теми нейронами слоя образцов, которые относятся к соответствующему классу;
3. Все веса слоя суммирования приравниваются к единицам;
4. В сети присутствует еще один слой, содержащий один нейрон, который определяет, какой из нейронов слоя суммирования дает максимальный выход, таким образом определяя класс входного вектора.

Как видно из первого пункта, недостатком данной сети является то, что для ее сохранения нужно большое количество памяти. Преимуществом является быстрота обучения, так как для определения весов последнего слоя здесь используется метод градиентного спуска.

Для решаемой в работе задачи сеть будет выглядеть так, как показано на рисунке 22, где N равно количеству примеров в обучающем наборе.

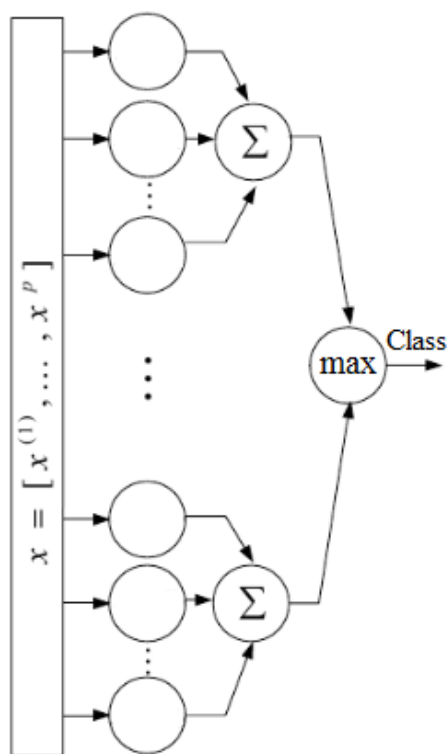


Рис. 21: Общая структура вероятностной нейронной сети [8].

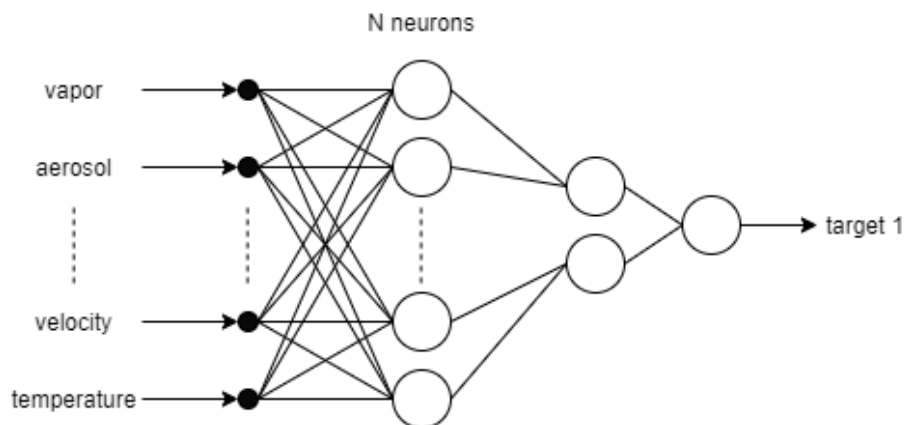


Рис. 22: Структура вероятностной нейронной сети, с помощью которой получена точность 90,4%.

Данная сеть была реализована на языке программирования Python с использованием библиотеки NeuPy. Точность классификации, полученная с использованием этой сети, составила 90,4%. Также было рассчитано значение среднеквадратической ошибки, которое составило 0,096.

Выводы

Были проведены исследования с использованием четырех типов нейронных сетей. Для каждого типа сети была найдена структура, дающая лучшие результаты, и получены значения точности и среднеквадратической ошибки классификации. Эти значения представлены в таблице 1.

Тип нейронной сети	Точность классификации, %	Ошибка классификации
Многослойный персептрон	89,1	0,082
Персептронный комплекс №1	90,4	0,081
Персептронный комплекс №2	91,6	0,078
Сеть радиально-базисных функций	91,6	0,069
Вероятностная нейронная сеть	90,4	0,096

Таблица 1: Значения точности и среднеквадратической ошибки, полученные с использованием различных типов нейронных сетей.

Самую низкую точность показал многослойный персептрон (рисунок 9). Было сделано предположение, что это связано с малым объемом обучающего набора. Для улучшения результата был применен один из методов, позволяющий решить проблему ограниченного объема данных, который состоит в использовании персептронного комплекса.

Было рассмотрено две структуры персептронных комплексов:

- без проскоков входа (рисунок 13),
- с двумя проскоками входа (рисунок 15).

Для первого комплекса набор признаков был разделен на группы, которые использовались в качестве входных переменных персептронов первого уровня таким образом, чтобы персептроны получали достаточно информации для независимого принятия решения о принадлежности исходного вектора к определенному классу. Это было сделано с использованием карты корреляции (рисунок 12). Как видно из таблицы 1, значения точности и ошибки улучшились. Проблемой этого комплекса является то, что

результатирующий персептрон определяет класс вектора, основываясь только на информации, полученной от персептронов первого уровня. Так что если персептроны первого уровня неправильно определяют класс вектора, результирующий персептрон, скорее всего, даст тот же результат.

Поэтому была рассмотрена вторая структура комплекса, в которой входами результирующего персептрона также являются два признака из исходного набора. Как видно из таблицы 1, такой подход позволил еще больше увеличить точность и уменьшить ошибку классификации.

Также были рассмотрены две нейронные сети, относящиеся к сетям радиально-базисного типа:

- сеть радиально-базисных функций (рисунок 19),
- вероятностная нейронная сеть (рисунок 22).

Из таблицы 1 видно, что из них лучший результат получится с использованием сети радиально-базисных функций. Это произошло потому, что, во-первых, в ней в качестве векторов-прототипов использовались центры кластеров, полученные с помощью метода k средних. То есть принадлежность входного вектора сети к какому-либо классу определялась с помощью его сравнения с вектором, в окрестности которого расположены некоторые векторы из обучающего набора, а не с самими векторами из обучающего набора. Соответственно сеть радиально-базисных функций в большей степени обладает способностью к обобщению, чем вероятностная нейронная сеть. Во-вторых, в той реализации сети радиально-базисных функций, которая использовалась в данной работе, значения векторов-прототипов корректировались во время обучения.

Из таблицы 1 видно, что наилучшие значения точности получились с использованием второго персептронного комплекса и сети радиально-базисных функций. Однако среднеквадратическая ошибка классификации, полученная с помощью сети радиально-базисных функций, получилось намного меньше. Также преимуществами сети РБФ перед персептронным комплексом являются простая структура и меньшее время обучения, так как в случае персептронного комплекса нужно последовательно провести обучение трех сетей с одним скрытым слоем.

Таким образом можно сделать вывод о том, что из рассмотренных типов нейронных сетей лучшим выбором для задачи прогнозирования опасно-

го конвективного явления (грозы) будет являться сеть радиально-базисных функций со структурой, показанной на рисунке 19, начальные значения векторов-прототипов в которой определяются с использованием метода k средних, а начальные значения коэффициентов β – по формуле (3).

Заключение

В данной работе проведено исследование применения искусственных нейронных сетей в задаче прогнозирования опасных конвективных явлений.

Была выбрана наиболее подходящая численная модель, позволяющая реализовать оперативный прогноз эволюции конвективного облака.

Написана программа, которая получает радиозондировки и преобразует их в формат входных данных численной модели. С использованием методов машинного обучения реализован алгоритм предварительной обработки выходных данных модели.

Прогнозирование опасных конвективных явлений (грозы) осуществлялось с использованием четырех типов нейронных сетей: многослойного персептрона, персептронного комплекса, сети радиально-базисных функций и вероятностной нейронной сети. Для каждого типа сети была найдена структура, с использованием которой получаются лучшие результаты. Был проведен анализ полученных результатов, на основе которого сделан вывод о том, что для осуществления прогнозирования грозы с использованием нейронной сети лучше всего использовать сеть радиально-базисных функций.

Список литературы

- [1] Назаренко, А. В. Опасные природные явления. Часть III. Опасные явления погоды конвективного происхождения / А. В. Назаренко. – Воронеж : Издательско-полиграфический центр Воронежского государственного университета, 2008.
- [2] Раба, Н. О. Программный комплекс компьютерного исследования атмосферных процессов для многоядерных процессоров: дис. ... канд. физ.-мат. наук – Санкт-Петербургский государственный университет, Санкт-Петербург, 2011 – 160 с.
- [3] P. Jeatrakul and K. W. Wong, “Comparing the Performance of Different Neural Networks for Binary Classification Problems,” The 8th International Symposium on Natural Language Processing, Bangkok, 20-22 October 2009, pp. 111-115. doi:10.1109/SNLP.2009.5340935
- [4] Солдатова, О. П. Нейроинформатика / О. П. Солдатова. – Самара, 2013.
- [5] Коробкова, С. В. Проблемы эффективной аппроксимации многомерных функций с помощью нейронных сетей / С.В. Коробкова // Известия ТРТУ. – 2006. – Т. 58, № 3. – С. 121-127.
- [6] Дударов С. П., Диев А. Н. Нейросетевое моделирование на основе перцептронных комплексов при малых объемах обучающих выборок // Математические методы в технике и технологиях – ММТТ-26: сб. трудов XXVI Междунар. науч. конф. — Секции 6, 7, 8, 9. — Нижегород. гос. техн. ун-т Нижний Новгород, 2013. — С. 114-116.
- [7] Моделирование свойств конструкционного композитного материала, армированного углеродными нанотрубками, с использованием перцептронных комплексов / С.В. Дударов [и др.] // Компьютерные исследования и моделирование. – 2015. – Т. 7, № 2. – С. 253-262.
- [8] Кветный, Р. Н. Вероятностные нейронные сети в задачах идентификации временных рядов / Р. Н. Кветный, В. В. Кабачий, О. О. Чумаченко

// Научные труды Винницкого национального технического университета. – 2010. – № 3.

- [9] Краткий словарь терминов по конвективным облакам [Электронный ресурс]. – Режим доступа: <http://meteoweb.ru/phen034.php>
- [10] Станкова Е.Н., Петров Д.А. Комплексная информационная система, предназначенная для формирования входных данных моделей конвективных облаков // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. Серия 10. 2015 Выпуск 3. Стр. 83-95.
- [11] Radial Basis Function Network (RBFN) Tutorial [Электронный ресурс]. – Режим доступа: <https://mccormickml.com/2013/08/15/radial-basis-function-network-rbf-tutorial/>
- [12] RBF layer for Keras [Электронный ресурс]. – Режим доступа: https://github.com/PetraVidnerova/rbf_keras