

MatLab с использованием выражений (1,2). На рис. 1 (b) в качестве примера представлены графики спектрального распределения естественного и искусственного излучений в относительных единицах при минимальном значении ϵ . Спектральный состав естественного излучения соответствовал дневному излучению с цветовой температурой 10000K [2]. Источниками искусственного излучения в рассматриваемом случае являлись группы белых и RGB светодиодов. Площадь под кривой спектрального распределения естественного излучения была условно принята за 1. Расчеты показали, что при минимальном значении ϵ , равном $4,6392 \cdot 10^{-7}$, соотношения произведений количества светодиодов и потребляемой ими мощности равны 0,038523, 0,098244, 0,014283, 0,42021 соответственно для синих, зеленых, красных светодиодов группы RGB и белого светодиода.

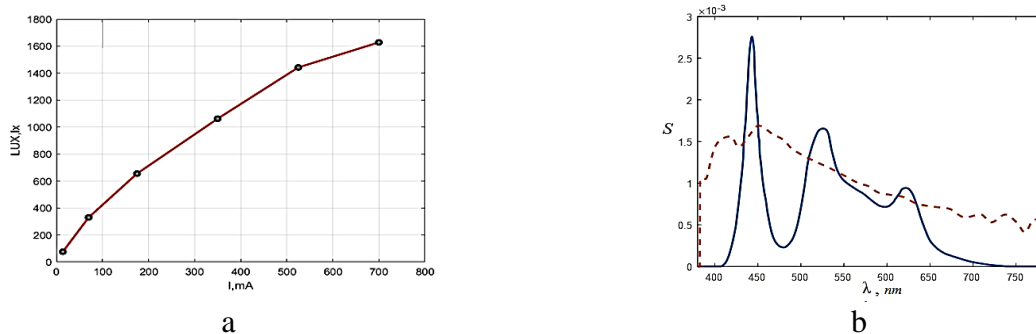


Рис.1. Зависимость освещенности LUX, лк, создаваемой светодиодной матрицей из 55 ортогонально расположенных белых светодиодов на расстоянии 350 мм от матрицы в центральной точке, от тока питания I, mA (a) и спектральное распределение естественного (пунктир) и искусственного (сплошная линия) излучений в относительных единицах при минимальном значении ϵ (b)

Сравнение графиков на рис.1(b) свидетельствует, что комбинация групп белого и RGB светодиодов является недостаточной для имитации спектра естественного излучения с их помощью вследствие наличия значительного пика в синей области и провалов в сине-зеленой и красной, т.е. для оптимизации необходимо дополнительно или альтернативно применять светодиоды с другим характером спектрального распределения.

Заключение

Разработана методика расчета параметров светодиодных источников искусственного излучения, которая позволяет анализировать максимально возможную степень совпадения их спектрального распределения со спектральным распределением естественного светового излучения.

Литература

1. Никифоров С. Температура в жизни и работе светодиодов. Ч.1 // Компоненты и технологии. 2005. - № 9. - с.48 - 54.
2. Blume C., Garbazza C., Spitschan M. Effects of light on human circadian rhythms, sleep and mood // *Somnologie (Berl)*. – 2019. - № 3(3). - p.147 – 156. Published online 2019 Aug 20. DOI: [10.1007/s11818-019-00215-x](https://doi.org/10.1007/s11818-019-00215-x)

ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ДАННЫХ И ИХ ПРОГНОЗИРОВАНИЕ

Ковшаров Г.Ю. (студент гр. ИТИ-22)

Гомельский государственный технический университет имени

П.О. Сухого, Гомель, Республика Беларусь

Научный руководитель – **Карась О. В.**

(ассистент кафедры «Информационные технологии» ГГТУ им. П.О. Сухого)

Аннотация: в работе проводится анализ взаимозависимостей параметров и прогнозирование данных.

Ключевые слова: прогнозирование данных, тепловая карта, уравнение регрессии.

Введение

Обработка данных и их прогнозирование предоставляет возможность предсказания каких-либо данных, на основе известных факторов. Пользой данной процедуры является предоставление возможности примерного предсказания какого-либо значения. Сфера применения методов прогнозирования охватывает области экономики и менеджмента, военных дисциплин и промышленности. В рамках данного доклада требуется сделать анализ, обработку и прогнозирование данных.

Результаты и обсуждение

Предположим, что дан набор со следующими параметрами: «Date», «Open», «High», «Low», «Close», «Volume», «Adj Close», и нужно спрогнозировать значение параметра «Close», в зависимости от значения зависимых параметров.

В первую очередь необходимо провести анализ зависимостей параметров друг от друга, это необходимо для нахождения значимых параметров. Сделать такой анализ можно при помощи различных способов: матрица диаграмм рассеивания, тепловая карта, матрица корреляции. Матрица диаграмм рассеивания – это инструмент визуализации данных, позволяющий сравнивать друг с другом несколько параметров, рассматривая их особенности и взаимосвязи; тепловая карта – это также инструмент, позволяющий визуализировать данные, а также их зависимости, при помощи различных цветов; матрица корреляции – это квадратная матрица, показывающая то, как данные коррелируют между собой. В данной задаче была использована тепловая картой изображённая на рисунке 1.

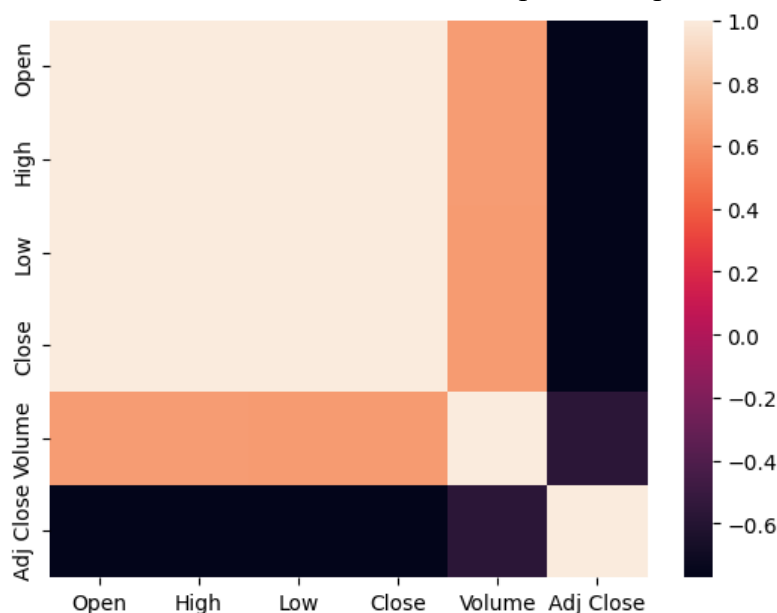


Рис 1. тепловая карта

Глядя на тепловую карту можно заметить, что на пересечении параметров расположены квадраты имеющие различные цвета, эти цвета определяют их зависимость: чем светлее квадрат, тем больше зависимость, например, параметр «Open» сильно зависит от параметра «Low», но никак не зависит от параметра «Adj Close».

Для определения зависимостей параметра «Close» нужно провести анализ тепловой карты. Глядя на неё, можно увидеть, что этот параметр сильно зависит от «Low», «High», «Open», но не зависит от параметра «Adj Close».

Теперь можно предположить, что для предсказания переменной, параметры зависят линейно и следует воспользоваться множественным линейным уравнением регрессии. Формула 1 – общая формула множественной линейной регрессии.

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n \quad 1$$

В данной формуле y – это предсказываемый параметр, b_i – коэффициенты регрессии для x_i , а x_i – значения зависимых параметров, которые нужно подставлять, для предсказания y .

Необходимо найти коэффициент регрессии, это можно сделать при помощи метода наименьших квадратов – один из методов теории ошибок, предназначенный для оценки неизвестных величин по результатам их измерений, содержащим случайные ошибки. Пример кода на *Python* реализующий поиск коэффициентов, при помощи библиотеки *numpy* представлен на рисунке 2.

```
X = data[['Days', 'Low', 'High', 'Open']].values
Y = data['Close'].values

# Вычисление коэффициентов множественной линейной регрессии
A = np.vstack([X.T, np.ones(len(X))]).T
coefficients = np.linalg.lstsq(A, Y, rcond=None)[0]

print(coefficients)

b1 = coefficients[0]
b2 = coefficients[1]
b3 = coefficients[2]
b4 = coefficients[3]

b0 = coefficients[4]
```

Рис 2. пример вычисления коэффициентов множественной линейной регрессии

Теперь, зная коэффициенты регрессии можно предсказывать значение параметра «Close» на определённую дату и с заданными параметрами «Low», «High», «Open», подставив их значения и коэффициенты в формулу 1.

Заключение

Результатом проделанной работы является нахождение коэффициентов регрессии, которые, в свою очередь, позволяют предсказывать параметр «Close» на определённую дату.

АЛГОРИТМЫ СРАВНЕНИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ

Кротов А.В. (студент гр. ИТИ-41)

*Гомельский государственный технический университет имени П.О.Сухого, Гомель,
Республика Беларусь*

Научный руководитель – **Курочка К.С.**

(к. т. н., заведующий кафедрой «Информационные технологии» ГГТУ им. П.О. Сухого)

Аннотация: исследуются современные методы сравнения текстов и анализируются их преимущества и недостатки.

Ключевые слова: алгоритмы сравнения текстов, обнаружение заимствований, внутренние заимствования.

Введение

Поиск плагиата среди текстовых документов является сложной, но в то же время востребованной задачей, особенно в академической и студенческой средах. Поиск плагиата – это комплекс средств, позволяющих определить степень схожести двух текстовых документов. Актуальность доклада обусловлена следующими факторами: 1) необходимость предотвращения копирования работ, которые были сделаны другими людьми; 2) активное интегрирование современных информационных технологий в образовательный процесс. Целью данного доклада является анализ существующих алгоритмов сравнения оригинального и проверяемого текстов для определения степени плагиата между ними и их применение для построения пользовательских приложений.

Результаты и обсуждение

Для сравнения текстов существует множество подходов, часть из которых основана на вычислении некоторого числа, описывающего степень схожести сравниваемого и оригинального текстов. Примером такого подхода является алгоритм вычисления косинусной меры подобия [1].

Для расчета косинусной меры подобия два текста разбиваются на отдельные слова, которые обрабатываются стандартными методами – приведение в нижний регистр, удаление