

Министерство науки и высшего образования Российской Федерации  
Московский Физико-технический институт  
(Государственный Университет)  
Физтех-школа прикладной математики и информатики  
Кафедра технологий цифровой трансформации

**Выпускная квалификационная работа**  
**"Развитие инструментов предиктивной аналитики в**  
**целях повышения эффективности мониторинга**  
**проектов в сфере жилищного строительства"**

**Студента 2-го курса Ефремова Сергея Владимировича**

**Научный руководитель**  
**кандидат экономических наук, доцент Помулев А. А.**

Москва, 2022

### **Аннотация**

Рассматривается задача улучшения инструментов предиктивной аналитики, использующихся при мониторинге проектов в сфере жилищного строительства. Исследованы предложенные ранее схемы решения этой проблемы, на основе изученных материалов разработан подход по улучшению оценки вероятности просрочки выплаты займа застройщиком на основании отчетности, публикуемой в открытом доступе и уровне зависимости от импортируемых комплектующих и материалов. Предложен, реализован и протестирован алгоритм, основанный на алгоритмах нейросетевого обучения с использованием чисел Шепли.

# Содержание

<b>1</b>	<b>Введение</b>	<b>4</b>
1.1	Цели и задачи работы . . . . .	6
<b>2</b>	<b>Постановка задачи</b>	<b>8</b>
2.1	Мониторинг проектов . . . . .	8
2.2	Эффективность мониторинга . . . . .	9
2.3	Предиктивная аналитика . . . . .	10
<b>3</b>	<b>Обзор действующей практики</b>	<b>11</b>
3.1	Анализ существующих проектов . . . . .	11
3.1.1	Проект 1 . . . . .	11
3.1.2	Проект 2 . . . . .	11
3.2	Текущее состояние финансирования в сфере жилищного строительства .	11
3.2.1	Ключевой топик 1 . . . . .	11
3.2.2	Ключевой топик 2 . . . . .	11
3.3	Процесс мониторинга проектов и методы оценки коммерческим банком .	11
3.4	Типы моделей предиктивной аналитики и их применение в кредитном процессе . . . . .	11
<b>4</b>	<b>Формальная постановка задачи</b>	<b>12</b>
4.1	Ключевые проблемы процесса мониторинга проектов . . . . .	12
4.2	Возможности внедрения с учетом консервативности и систем безопасно- сти банка . . . . .	12
<b>5</b>	<b>Описание модели</b>	<b>13</b>
5.1	Этап предобработки данных . . . . .	13
5.2	Ядро модели . . . . .	13
<b>6</b>	<b>Результаты работы алгоритма</b>	<b>14</b>
6.1	Пример полученных результатов - ключевые атрибуты . . . . .	14
6.2	Сравнение результатов с другими методами . . . . .	14
6.3	Сравнение результатов с оценкой предложенной метрики качества . . . .	14
<b>7</b>	<b>Экономический эффект от внедрения модели</b>	<b>15</b>
<b>8</b>	<b>Заключение</b>	<b>16</b>

# 1 Введение

На сегодняшний день можно с уверенностью говорить о том, что биометрические технологии идентификации вошли в нашу жизнь и используются во всех сферах человеческой деятельности, в том числе и в системах контроля. Существует целый ряд биометрических модальностей, позволяющих идентифицировать человека: рисунок папиллярных линий пальца, радужка глаза, походка, голос, почерк. Такое разнообразие методов объясняется тем, что при богатстве потенциальных сценариев использования биометрических модальностей, каждая из них может быть использована по-разному. Применение биометрических технологий уже активно осуществляется при проведении платежных операций в банковской сфере, особо нуждающейся в защите данных. Здесь они заменяют уже ставшие традиционными методы идентификации пользователей: ПИН-коды, пароли и т.д. Биометрические технологии имеют широкую перспективу использования, так как обладают целым рядом неоспоримых преимуществ, среди которых можно выделить неотчуждаемость. Например, пароль может быть передан или украден третьим лицом, и система все равно идентифицирует его как исходного пользователя, тогда как биометрическими модальностями труднее воспользоваться другому человеку или украсть без ведома обладателя.

Использование биометрических систем настолько повсеместно, что обыватель даже не замечает, как часто и повсеместно ими пользуется. Так, для разблокировки мобильных устройств, применяется отпечаток пальца, снимок лица или радужки. Еще одним преимуществом биометрических систем считается удобство взаимодействия с ними пользователя. Их эффективность работы основывается на способности обработки поступающих данных с учетом изменчивости состояния и условий окружающей среды. То есть, даже в случае значительных искажений параметров входных данных, в связи с изменчивостью состояний (например, ухудшение условий освещенности, затемнение ресницами, отвод взгляда), системы биометрической идентификации, использующие изображения глаза, должны выдавать корректный результат и аутентифицировать пользователя адекватно. Также неоспоримо важным фактором для эффективности систем такого рода, является необходимость работы в условиях ограниченности не только вычислительных ресурсов, но и времени, требующегося на распознавание.

Благодаря стремительному развитию технологий, с каждым днём все больше расширяется сфера практического применения технических средств цифровой обработки изображений. Одним из направлений стало распознавание человека на основе изображения глаза (рис. 1). Эта идея не нова, ещё в 1936 году офтальмологи [18] предложили использовать радужную оболочку глаза для распознавания. Далее исследователями

всего мира разрабатывались различные методы распознавания человека по его радужной оболочке, но по-настоящему сильный рост интереса произошел только в начале XXI века. Это объясняется качественными изменениями условий жизни общества: эволюция средств регистрирования и обработки изображений, необходимость распознавания личности в связи с террористической угрозой [9, 13, 14]. Наряду с идентификацией обработка изображения радужки применяется в медицинских целях, где также важно выделение области интереса.

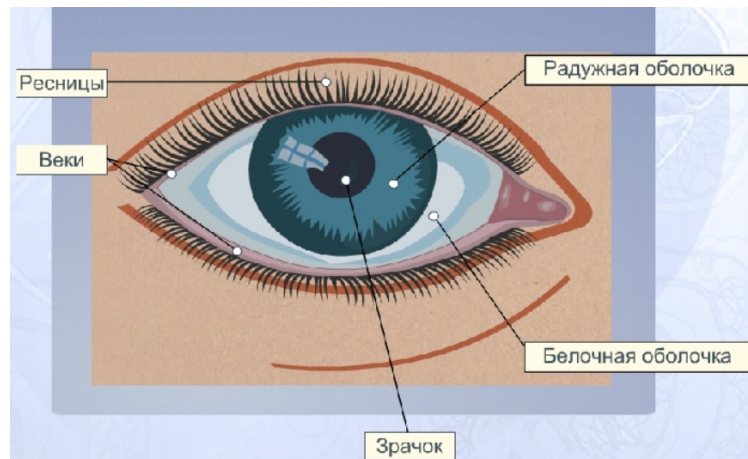


Рис. 1: Внешнее строение глаза.

Однако получить качественное и обособленное изображение радужки на практике не так просто, и важной задачей оказывается необходимость корректно определять границы глаза, задаваемые веками. Исследования в этом направлении породили различные модели, которые также используют в биометрических системах распознавания, поскольку уточнение границ и исключение нерелевантных областей оказывает значительное влияние на точность результата.

В ходе данного исследования разработан метод выделения века как параметрической кривой на изображении глаза. Он состоит в определении параметров кривой максимизацией суммы вертикального градиента яркости, рассматриваемого в окрестности века.

## 1.1 Цели и задачи работы

**Цель и задачи исследования.** Целью исследования является построение модели предиктивной аналитики, которая позволит повысить эффективность процесса мониторинга проектов коммерческим банком в сфере жилищного строительства и улучшить качество прогнозирования вероятности просрочки платежа по сравнению с существующими моделями. Для реализации этой цели были поставлены следующие задачи:

- изучить определение понятий: «мониторинг», «эффективность мониторинга», «предиктивная аналитика» для использования в настоящем исследовании;
- провести анализ существующих проектов и динамики их развития в сфере жилищного строительства;
- ознакомиться с текущим состоянием финансирования проектов в сфере жилищного строительства и нормативно-правовой базой;
- изучить процесс мониторинга проектов и методы их оценки коммерческим банком;
- исследовать типы моделей предиктивной аналитики и их применение в кредитном процессе;
- выделить основные проблемы процесса мониторинга проектов и определить возможности их решения с использованием инструментария предиктивной модели
- разработать алгоритм внедрения разработанного инструментария в бизнес-процесс мониторинга
- рассчитать экономический эффект от внедрения модели

**Научная новизна.** Используется нейросетевой подход к определению вероятности банкротства заемщика с выделением признаков, вносящих максимальный вклад с помощью, чисел Шепли. В работе предлагается коэффициент, позволяющий оценить зависимость застройщика от импортных комплектующих и материалов, а также уровень потенциального риска, обусловленного политическими ограничениями.

**Методы исследования.** Алгоритмы реализованы на языке программирования Python с использованием библиотек `|||`.

**Практическая ценность.** Полученная модель может быть использована в качестве встраиваемого модуля. Например, с её помощью можно:

- корректировать оценку вероятности просрочки платежа застройщиком, учитывая его зависимость от импортируемых компонентов;
- дополнять существующие системы мониторинга объектов строительства показателем уровня зависимости от импортных компонентов и моделью оценки наиболее важных показателей, влияющих на просрочку.

## 2 Постановка задачи

### 2.1 Мониторинг проектов

изучить определение понятий: «мониторинг», «эффективность мониторинга», «предиктивная аналитика» для использования в настоящем исследовании

Мониторинг проекта - процесс измерения показателей выполнения проекта, сбора данных об исполнении проекта, информационного обслуживания управления проектом с целью выявления его соответствия желаемому результату и плану, с последующим представлением и распространением полученных данных.

Под контролем проекта понимается процесс сравнения фактических значений контрольных показателей с запланированными, последующего анализа отклонений, оценки тенденций и прогнозирования возможных альтернатив, разработки корректировок хода реализации проекта для улучшения прогноза.

Основными целями контроля и мониторинга инвестиционных проектов можно считать обеспечение:

- своевременного достижения целей проекта с учетом согласованной стоимости;
- срочности, возвратности, платности и целевого использования предоставляемых банком кредитных ресурсов для финансирования проекта;
- своевременного информирования руководства банка о выявленных проблемах, прогнозирования рисков реализации проекта и разработка мер по их снижению;
- достижения заложенных в проекте показателей социально-экономической эффективности.

Чаще всего при реализации проектов в сфере жилищного строительства выделяют следующие виды мониторинга:

- мониторинг хода реализации инвестиционного проекта (сроков выполнения работ, бюджета проекта, расчетного времени окончания работ и расчетной стоимости проекта, организация технадзора и контрольных проверок);
- финансовый мониторинг (финансово-экономического состояния заемщика, исполнителя проекта, поручителей, гарантов, обеспечения по кредиту/кредитной линии; денежного потока, коэффициентов покрытия, целевого использования средств, исполнения заемщиком обязательств перед банком);
- мониторинг эффективности инвестиционного проекта (показателей, которые предусмотрены положением об экспертизе проектов банка).



Такое разделение обуславливается необходимостью не только контролировать текущую операционную деятельность, ведущуюся по проекту, исполнение финансовых обязательств участниками проекта и целевое использование средств, но и конечные результаты этой деятельности, которые выражаются в достижении целей проекта и достигнутой социально-экономической значимости.

Основными элементами систем мониторинга инвестиционных проектов являются:

- финансовая, техническая и иная отчетность заемщика;
- экспертные оценки банковских специалистов по направлениям реализации проекта, независимые эксперты (технический надзор, финансовый аудит);
- календарно-сетевые графики работ, расчеты сроков ввода объекта в эксплуатацию и суммарной стоимости работ;
- данные автоматизированных информационных систем мониторинга инвестиционных проектов.

Последние и будут рассмотрены в первую очередь в данной работе.

Ключевые этапы мониторинга проектов:

- этап подготовки проекта (начинается с момента одобрения займа/кредитной линии и заканчивается выделением финансовых средств);
- инвестиционная стадия проекта (непосредственное финансирование проекта);
- этап эксплуатации (следует до полного исполнения заемщиком платежных обязательств перед банком).

## 2.2 Эффективность мониторинга

Построением эффективных систем мониторинга занимались многие исследователи Д. Боуэр, Дж.Филлипс, Р.Фартел, Х.Керцнер[здесь будут ссылки на литературу]. Мониторинг в современных реалиях представляет из себя комплексную функцию проектного управления, в которую входит процедуры сбора, анализа и передачи информации о ходе реализации проекта, которая позволяет решить проблему своевременного принятия решений по проекту.

Основные задачи, которые решают системы мониторинга:

- определение совокупности отслеживаемых индикаторов;
- организация обработки и агрегирования полученной информации;

- генерация текущей отчетности по проекту;
- интеграция функции мониторинга в информационную архитектуру предприятия, реализующего проект.

Принятие управленческих решений о формировании и развитии системы мониторинга проектов, о требуемом кадровом, техническом и финансовом обеспечении неизбежно связано с дополнительными затратами. Однако, потенциальные угрозы от финансирования убыточных или высокорисковых проектов также способны привести к значительным издержкам. Все это остро ставит вопрос о необходимости эффективного мониторинга проектов.

Основными подходами к изучению эффективности проектов являются:

- целевой (предполагает анализ степени достижения целевых значений показателей);
- динамический (учитывает скорость изменения исследуемых показателей во времени и относительно друг друга);
- затратный (основан на сопоставлении затрат и результатов);
- ресурсный (исследует степень рациональности расходования ресурсов).

## 2.3 Предиктивная аналитика

Предиктивной аналитикой или продвинутой аналитикой называют ряд аналитических и статистических методов прогнозирования действий и поведения в будущем. В основе лежат статистические модели, позволяющие находить закономерности в исторических и транзакционных данных, что позволяет выделять потенциальные риски и возможности. Ключевые этапы составляющие процесс предиктивного анализа: подключение к данным, анализ и визуализация результатов исследований, развитие предложений и моделей данных, применение предиктивных моделей, оценка и прогнозирование будущих результатов.

В основе предиктивной аналитики лежит выявление связей между данными историческими и прогнозными результатами на их основе. Верхнеуровнево алгоритмы предиктивного анализа можно разделить на контролируемое и неконтролируемое обучение.

Контролируемое обучение принято разделять на две ключевые категории: регрессию для количественных ответов и классификацию для определения фактической принадлежности ответа к той или иной группе.

Неконтролируемое обучение применяется для получения выводов из входных данных без разметки. Наиболее распространенный вид такого анализа - кластеризация, которую используют для поиска скрытых закономерностей в данных.

### **3 Обзор действующей практики**

#### **3.1 Анализ существующих проектов**

##### **3.1.1 Проект 1**

##### **3.1.2 Проект 2**

#### **3.2 Текущее состояние финансирования в сфере жилищного строительства**

##### **3.2.1 Ключевой топик 1**

##### **3.2.2 Ключевой топик 2**

#### **3.3 Процесс мониторинга проектов и методы оценки коммерческим банком**

#### **3.4 Типы моделей предиктивной аналитики и их применение в кредитном процессе**

## **4 Формальная постановка задачи**

### **4.1 Ключевые проблемы процесса мониторинга проектов**

### **4.2 Возможности внедрения с учетом консервативности и систем безопасности банка**

## 5 Описание модели

### 5.1 Этап предобработки данных

### 5.2 Ядро модели

## **6 Результаты работы алгоритма**

**6.1 Пример полученных результатов - ключевые атрибуты**

**6.2 Сравнение результатов с другими методами**

**6.3 Сравнение результатов с оценкой предложенной метрики качества**

## **7 Экономический эффект от внедрения модели**



## 8 Заключение

## Список литературы

- [1] *Adam M., Rossant F. et al.* Eyelid localization for iris identification // *Radioengineering.* — 2008. — Vol. 17, no. 4. — Pp. 82–85.
- [2] *Adam M., Rossant F. et al.* Reliable eyelid localization for iris identification // *Advanced Concepts for Intelligent Vision Systems Conference Proceedings.* — 2008. — Pp. 1062–1070.
- [3] *Ballard D.* Generalizing the hough transform to detect arbitrary shapes // *Pattern Recognition.* — 1981. — Vol. 13, no. 2. — Pp. 111–122.
- [4] *Canny J.* A computational approach to edge detection // *IEEE Trans. Pattern Anal. Mach. Intell.* — 1986. — Vol. 8, no. 6. — Pp. 679–698.
- [5] *Daugman J.* High confidence visual recognition of persons by a test of statistical independence // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* — 1993. — Vol. 15, no. 11. — Pp. 1148–1161.
- [6] *Daugman J.* The importance of being random: statistical principles of iris recognition // *Pattern Recognit.* — 2003. — Vol. 36. — Pp. 279–291.
- [7] *Daugman J.* How iris recognition works // *IEEE Trans. Circuits Syst. Video Techn.* — 2004. — Vol. 14. — Pp. 21–30.
- [8] *Deans S.* The radon transform and some of its applications // *New York: John Wiley and Sons.* — 1983.
- [9] Future challenges based on the multiple biometric grand challenge // NIST information access division: Multiple biometric grand challenge. — 2010.
- [10] *Kang B., Park K.* A robust eyelash detection based on iris focus assessment // *Pattern Recognition Letters.* — 2007. — Vol. 28, no. 13. — Pp. 1630–1639.
- [11] *Masek L.* Recognition of human iris patterns for biometric identification // *Measurement.* — 2003. — Vol. 32, no. 8. — Pp. 1502–1516.
- [12] *Min T., Park R.* Comparison of eyelid and eyelash detection algorithms for performance improvement of iris recognition // *Conference (International) on Image Processings Proceedings.* — 2008. — Pp. 257–260.
- [13] Multiple biometric evaluation // <http://www.nist.gov/itl/iad/ig/mbe.cfm>. — 2009.

- [14] Multiple biometric grand challenge // <http://www.nist.gov/itl/iad/ig/mbgc.cfm>. — 2007.
- [15] Wildes R. Iris recognition: an emerging biometric technology // *Proceedings of the IEEE*. — 1997. — Vol. 85, no. 9. — Pp. 1348–1363.
- [16] Xiangde Z., Qi W. et al. Noise detection of iris image based on texture analysis // *Chinese Control and Decision Conference Proceedings*. — 2009. — Pp. 2366–2370.
- [17] Yang L., Wu T. et al. Eyelid localization using asymmetric canny operator // *Conference (International) on Computer Design and Applications Proceedings*. — 2010. — Pp. 533–535.
- [18] Вельховер Е.С., Шульпина Н.Б., Алиева З.А. и др. Иридодиагностика // М.: Медицина. — 1988. — Р. 240.
- [19] Гонсалес Р., Вудс Р. Цифровая обработка изображений // *Техносфера*. — 2012.
- [20] Дьяконов В.П. Справочник по алгоритмам и программам на языке бейсик для персональных ЭВМ // М.: Наука. — 1987.
- [21] Смирнов Д.А., Матвеев И.А. Определение границ век на изображении глаза методом активных контуров // *Труды ИСА РАН, Динамика неоднородных систем*. — 2006. — Vol. 25. — Pp. 200–207.
- [22] Соломатин И.А., Матвеев И.А., Новик В.П. Определение видимой области радужки классификатором текстур с опорным множеством // *Автоматика и телемеханика*. — 2018. — Pp. 127–143.