



## **Программный продукт**

### **«Голосовой скоринг»**

Документация, содержащая описание функциональных характеристик программного обеспечения и информацию, необходимую для установки и эксплуатации программного обеспечения



## Оглавление

1 Назначение Системы.....	3
2 Функциональные возможности Системы .....	5
3 Информация, необходимая для эксплуатации Системы.....	7
4 Установка Системы.....	9
4.1 Порядок установки и настройки Системы в ручном режиме .....	9
4.1.1 Загрузка образов в Docker.....	9
4.1.2 Запуск контейнеров с сервисами.....	9
4.2 Установка Системы в автоматизированном режиме .....	10
4.2.1 Порядок установки в автоматизированном режиме .....	10
4.2.2 Заполнение файла production.yml.....	10
4.2.3 Запуск Ansible-скриптов .....	10



## 1 Назначение Системы

Программное обеспечение «Голосовой скоринг» (далее – Система) состоит из сервисов, позволяющих работать с различными моделями. Состав моделей и сервисов может отличаться в зависимости от решаемых задач.

Система позволяет анализировать речь клиента и оператора с целью выявления паттернов, относящихся к спектру бизнес-задач. Входными данными для Системы являются текстовые данные и другие характеристики речи. Выходными данными моделей могут быть скорбалл или иной классифицирующий признак.

Пользователь с помощью API отправляет в Систему запрос на скоринг физического лица, Система проводит скоринг в соответствии с заранее определенными параметрами и с помощью API отправляет ответ со скорбаллом пользователю Системы.

Система помогает повысить качество оценки высказываний клиентов и операторов с помощью анализа спонтанной речи абонента при его устном ответе на предварительно подготовленные вопросы. Целевые модели могут использовать как доступные статические факторы, накопленные, так и информацию, извлекаемую из голосового канала при разговоре (как текстовую – обработка естественного языка, так и невербальную – определение эмоций, интонаций, правдивости ответов).

Работа Системы включает следующие этапы:

1. Консолидация источников данных (Data Mining).
2. Определение целевой переменной.
3. Разработка модели выявления критериев эффективности.
4. Анализ транскрибации звонков (Natural Language Processing).
5. Определение эмоций.

В качестве источника для извлечения данных из речи клиентов могут служить автоматизированная система заказчика или другие внешние ресурсы. В разных областях деятельности извлечение данных из речи клиентов и их обработка позволит:

1. Уменьшить фрод и другие риски при работе с клиентами.
2. Быстро найти качественных кандидатов с необходимыми компетенциями на массовые вакансии.
3. Повысить качество скоринговой оценки кредитного портфеля или отдельных заемщиков.



После анализа диалога клиента и оператора полученная скоринговая оценка фиксируются в операционной базе данных (или в автоматизированной системе заказчика) для автоматического определения дальнейшего сценария развития бизнес-процессов.

Система применяется в сферах деятельности, где требуется дистанционная коммуникация: финансовая отрасль, страхование, взыскание задолженности, подбор персонала, сферы обслуживания и продаж и др.



## 2 Функциональные возможности Системы

Система предоставляет следующие функциональные возможности:

1. Оценка высказываний клиента или оператора в зависимости от бизнес-требований.
2. Предсказание поведения клиента по распознанной информации из голосового канала и историческим данным в режиме реального времени.
3. Предсказание необходимости воздействия на клиента по историческим данным.
4. Поддержка различных моделей в зависимости от бизнес-требований.
5. Сигнализация оценки на стороне оператора.
6. Распознавание эмоций оператора и абонента (опционально).
7. Распознавание речи оператора и абонента (опционально).

Примеры сервисов и моделей, который могут быть включены в поставку:

### 1. Модель для прогнозирования воздействия (Проблемная задолженность).

Цель использования модели – определить эффективное воздействие по взысканию проблемной задолженности для каждого клиента.

Оператор задает клиенту вопросы по заранее согласованному скрипту. Клиент отвечает на вопросы оператора. Аудио-ответы клиента обрабатываются онлайн системой оптимизации взыскания на базе моделей Голосового скоринга. Формируется перечень наиболее эффективных следующих шагов по взысканию с оценкой по каждому клиенту и влиянию на портфель в целом.

Точность модели (коэффициент Джини) – 72%.

Результат после применения модели – в 2 раза сокращены издержки во взыскании.

### 2. Модель для предсказания жалоб клиентов контакт-центра.

Цель использования модели – сократить количество жалоб от клиентов.

Каждая коммуникация с клиентом анализируется в онлайн-режиме, и оператору отображается вероятность жалобы клиента.

Сокращение количества жалоб достигается с применением анализа текстов диалогов с помощью набора моделей Голосового скоринга, моделей детектирования эмоций (негатив, нейтральная, позитив) и ad hoc сценариев для снятия напряжения.

Точность модели (коэффициент Джини) – 91%.

Результат после применения модели – на 10% снижен поток жалоб.

### 3. Модель для подбора на массовые должности для HR-службы.



Цель использования модели – осуществить релевантный набор сотрудников на массовые должности.

В процессе собеседования голосовой бот проводит интервью, а кандидат отвечает на вопросы. Аудио-ответы кандидата обрабатываются моделями Голосового скоринга. Сервис осуществляет анализ интервью по перечню вопросов-ответов. Модель ранжирует кандидатов на перспективных и неперспективных на основе их ответов. В результате рекрутер получает отчет о кандидатах с их оценками.

Точность модели (коэффициент Джини) – 36%.

Результат после применения модели – релевантный отбор кандидатов, большой охват кандидатов, принятие итогового решения остается за рекрутером.

#### **4. Фрод-аналитика по кредитам физических лиц.**

Цель использования модели – уменьшить проблемную задолженность в кредитном портфеле на этапе его формирования.

Андеррайтер задает вопросы по сценарию андеррайтинга, а клиент отвечает на вопросы. Аудио-ответы кандидата обрабатываются моделями Голосового скоринга. Сервис осуществляет анализ звонков андеррайтеров и отмечает потенциально проблемных клиентов. В результате андеррайтер получает отчет о клиентах с их оценками.

Точность модели (коэффициент Джини) – 56%.

Результат после применения модели – в 10% звонков выявляются 50% фрода.

#### **5. Повышение эффективности follow-up в телемаркетинге.**

Цель использования модели – повысить конверсию follow-up звонков.

Сервис осуществляет анализ диалога при первом обзвоне роботом и выделяет группы клиентов с наибольшей готовностью приобрести продукт для проведения follow-up звонков. Повышение конверсии follow-up звонков достигается моделью, оценивающей «правдивость» намерения клиента оформить продукт.

Балл модели, оценивающей речь клиента, который не оформил продукт, варьируется от 0 до 40%. Балл модели, оценивающей речь клиента, который оформил продукт, варьируется от 40 до 100%.

Точность модели (коэффициент Джини) – 43%.

Результат после применения модели – экономия на follow-up звонках составляет 3-5%.



### **3 Информация, необходимая для эксплуатации Системы**

Для корректного функционирования Системы требуется выполнение минимальных аппаратных и программных требований к составным компонентам, в зависимости от набора и объема бизнес-задач требования могут значительно отличаться. В [таблице 1](#) приведены требования для минимального функционирования системы, с расчетом на 100 операторов с необходимостью распознавать эмоции и речь.

Система может линейно масштабироваться в зависимости от изменения бизнес-требований и нагрузки. Масштабирование осуществляется за счет увеличения числа нод с сервисами.

Рекомендовано не допускать нагрузку на серверах выше 50%, для осуществления требований к отказоустойчивости, а также для обеспечения возможности модернизации и изменения конфигурации Системы без остановки сервиса.

Рекомендовано использование автоматизированных систем для доставки (CI/CD, например, Jenkins\Gitlab) Системы до конечных серверов, чтобы исключить человеческий фактор из процесса установки, обновления и внесения изменения в конфигурационные файлы Системы.

Рекомендовано использование любых принятых систем мониторинга работы Системы и операционных систем (Zabbix, Nagios и т. д.) для своевременного обнаружения проблемы в работе оборудования, ОС, Системы. По запросу в комплект поставки могут быть включены шаблоны для Zabbix.

Требований к резервному копированию не предусмотрено. Результаты работы Системы хранятся в БД на стороне заказчика, которые резервируется в соответствии с регламентом.

Рекомендовано хранить логи системы как минимум 7 дней и использовать сторонние системы хранения логов (например, ELK) для изучения в случае аварии.



Таблица 1 – Минимальные требования, предъявляемые к компонентам Системы

Роль	CPU	RAM (ГБ)	HDD/СХД (ГБ)	Количество серверов		ОС	ПО
				Минимальная конфигурация	Конфигурация с отказоустойчивостью		
<b>Основные компоненты</b>							
Сервер сервисов ПО «Голосовой скоринг»	16	116	200	1	2	RHEL 7.7, CentOS 7.7	Рекомендовано использование GPU Nvidia с поддержкой CUDA для наиболее эффективной работы
<b>Оptionальные компоненты</b>							
Сервер распознавания речи	8	16	50	8	12	RHEL 7.7, CentOS 7.7	
Сервер распознавания эмоций	8	8	50	2	3	RHEL 7.7, CentOS 7.7	





## 4 Установка Системы

Установка Системы возможна в ручном или в автоматизированном режимах с использованием Ansible-скриптов из комплекта поставки. Рекомендуется использование автоматизированного варианта.

### 4.1 Порядок установки и настройки Системы в ручном режиме

1. Загрузка образов в Docker.
2. Запуск контейнеров с сервисами.

#### 4.1.1 Загрузка образов в Docker

В зависимости от поставки, комплект будет содержать один или несколько docker-образов, сохраненных в формате tar.gz.

Для загрузки docker-образа с сервисами в систему, скопируйте docker-образ и выполните команду:

```
docker load < gs-services.tar.gz
```

где *gs-services.tar.gz* – имя образа из контейнера.

#### 4.1.2 Запуск контейнеров с сервисами

В зависимости от комплекта поставки может быть разное число сервисов и, соответственно, контейнеров. Подробная информация о необходимых сервисах содержится в поставке.

Запустите контейнер:

```
docker run -d --restart unless-stopped -network=host --hostname abc-gs-services -- name abc-gs-services nexus.activebc.ru:5000/abc-gs-services-имя-сервиса
```

Для некоторых сервисов требуются конфигурационные файлы. Эта информация содержится в поставке. В этом случае необходимо сформировать конфигурационный файл и примонтировать его в контейнер по пути */config.json* при запуске контейнера. Например:

```
docker run -d --restart unless-stopped -network=host --hostname abc-gs-services -- name abc-gs-services -v /config.json:/config.json nexus.activebc.ru:5000/abc-gs-services-имя-сервиса
```



---

Проверьте, что контейнеры запустились, используя команду

*docker ps*

## 4.2 Установка Системы в автоматизированном режиме

В комплект поставки входят Ansible-скрипты, которые содержат все необходимые роли для установки поставки, а также playbook, который содержит последовательность вызова ролей для данной поставки.

### 4.2.1 Порядок установки в автоматизированном режиме

1. Заполнение файла production.yml.
2. Запуск Ansible-скриптов.

### 4.2.2 Заполнение файла production.yml

Для выполнения Ansible-скрипта необходимо заполнить файл инфраструктуры (production.yml), который будет содержать информацию о серверах, на которые будет производиться установка сервисов, а также базовая информация, необходимая для формирования конфигурационных файлов из шаблонов. Для заполнения рекомендуется использовать текстовый редактор «Vim» или «Notepad++».

### 4.2.3 Запуск Ansible-скриптов

Рекомендовано использование Ansible не ниже версии 2.5.1. После заполнения production-файла запустите playbook командой:

*ansible-playbook -i production.yml abc-gs-playbook.yml*

Ansible последовательно загрузит docker-образы на сервера и запустит необходимые сервисы. Также в Ansible включены базовые тесты системы.