

ТЕХНИЧЕСКАЯ СПЕЦИФИКАЦИЯ И КАЛЕНДАРНЫЙ ПЛАН РАБОТ

По договору № _____ от _____ 2018 года

1. Некоммерческое акционерное общество «Казахский национальный исследовательский технический университет имени К.И. Сатпаева»

1.1 По приоритету: 3. Информационные, телекоммуникационные и космические технологии, научные исследования в области естественных наук

1.2 По подприоритету: 3.1 Интеллектуальные информационные технологии. Машинное обучение (machine learning)

1.3 По теме проекта: № AP05135175 «Разработка и внедрение системы по распознаванию рукописных адресов письменной корреспонденции АО «Казпочта» с использованием машинного обучения»

1.4 Общая сумма проекта 30 000 000 (тридцать миллионов) тенге, в том числе с разбивкой по годам, для выполнения работ согласно пункту 3:

- на 2018 год - в сумме 10 000 000 (десять миллионов) тенге;
- на 2019 год - в сумме 10 000 000 (десять миллионов) тенге;
- на 2020 год - в сумме 10 000 000 (десять миллионов) тенге.

2. Характеристика научно-технической продукции по квалификационным признакам и экономические показатели

2.1 Направление работы: Прикладные исследования в области распознавания текста.

2.2 Область применения: Распознавание рукописного текста

2.3 Конечный результат:

- за 2018 год: Набор существующих алгоритмов и сбор датасетов. Архитектура разрабатываемой системы;

- за 2019 год: Модель оптического распознавания рукописного текста;

- за 2020 год: Программный комплекс для распознавания рукописного текста. За весь период реализации проекта будут опубликованы 2 статьи в рецензируемых зарубежных научных изданиях (Scopus), в одном или нескольких из следующего списка: «Applied Artificial Intelligence», «Statistics, Optimization and Information Computing», «Journal of Artificial Intelligence», «Information and Control», «International Journal of Computing Science and Mathematics», «Applied Mathematics and Information Sciences» и 2 статьи в отечественных научных журналах с ненулевым импакт-фактором..

2.4 Патентоспособность: непатентоспособен.

2.5 Научно-технический уровень (новизна): Реализация данного проекта позволит разработать систему автоматического распознавания адреса почтовой корреспонденции, включая рукописные адреса, максимально приближенной к имеющимся бизнес-процессам в курьерских службах. Автоматизация данного бизнес-процесса по регистрации почтового отправления существенно снизит издержки курьерских служб по пересылке корреспонденции и снизит время ее доставки.

2.6 Использование научно-технической продукции осуществляется: Заказчиком и Исполнителем совместно.

2.7 Вид использования результата научной и (или) научно-технической деятельности: Публикации.

3. Наименование работ, сроки их реализации и результаты

Шифр задания, этапа	Наименование работ по Договору и основные этапы его выполнения*	Срок выполнения*		Ожидаемый результат*
		начало	окончание	
1	Анализ существующих алгоритмов распознавания рукописного текста, сбор данных и проектирование разрабатываемой системы	Январь 2018	1 ноября 2018	Будет проведен анализ существующих алгоритмов распознавания рукописного текста, сбор данных и проектирование разрабатываемой системы. Будет получен набор существующих алгоритмов и собраны датасеты. Будет разработана архитектура системы
1.1	Обзор существующих решений с использованием и без машинного обучения для задачи оптического распознавания текста	Январь 2018	Март 2018	Будет проведен обзор существующих решений с использованием и без машинного обучения для задачи оптического распознавания текста. Будет изучен опыт зарубежных исследований в области распознавания рукописного текста
1.2	Анализ существующих решений	Апрель 2018	Июнь 2018	Будет проведен анализ существующих решений. Будут выявлены потенциально сильные и слабые стороны, используемые на данный момент алгоритмов
1.3	Сбор данных - набор датасетов из существующих реальных данных (собрать множество фотографий рукописного текста, однозначно отображаемое на множество цифрового текста)	Июль 2018	Сентябрь 2018	Будет произведен сбор данных - набор датасетов из существующих реальных данных (собрать множество фотографий рукописного текста, однозначно отображаемое на множество цифрового текста) и набор датасетов из существующих реальных релевантных данных.
1.4	Построение архитектуры программного решения	Октябрь 2018	1 ноября 2018	Будет построена архитектура программного решения. Будет разработана архитектура системы
2	Разработка собственной модели оптического распознавания рукописного текста и его отладка	Январь 2019	1 ноября 2019	Будет разработана собственная модель оптического распознавания рукописного текста и его отладка.
2.1	Опыты с существующими решениями на собранном датасете и сравнение полученных результатов для разных решений	Январь 2019	Март 2019	Будут проведены опыты с существующими решениями на собранном датасете и сравнение полученных результатов для разных решений. Будет проведен сравнительный анализ для разных существующих решений на собранном датасете
2.2	Построение нескольких методов машинного обучения для решения задачи и проведение опытов с каждым решением	Апрель 2019	Июнь 2019	Будут построено несколько методов машинного обучения для решения задачи и проведение опытов с каждым решением. Будет выбран наилучший метод для решения задачи

2.3	Построение окончательной модели выбранного решения	Июль 2019	Сентябрь 2019	Будет построена окончательная модель выбранного решения. Будет построена модель оптического распознавания рукописного текста
2.4	Тестирование и отладка разработанной модели	Октябрь 2019	1 ноября 2019	Будет проведено тестирование и отладка разработанной модели. Будет получена модель оптического распознавания рукописного текста с улучшениями
3	Построение программного обеспечения	Январь 2020	1 ноября 2020	Будет построено программное обеспечение для распознавания рукописного текста
3.1	Сравнение с результатами показанными уже существующими решениями	Январь 2020	Март 2020	Будет проведено сравнение с результатами показанными уже существующими решениями
3.2	Разработка программного продукта	Апрель 2020	Июнь 2020	Будет разработан программный продукт – программное обеспечение.
3.3	Тестирование полученной системы оптического распознавания рукописного текста	Июль 2020	Сентябрь 2020	Будет проведено тестирование полученной системы оптического распознавания рукописного текста. Будет получено готовое решение для распознавания рукописного текста
3.4	Написание итогового отчета, публикаций	Октябрь 2020	1 ноября 2020	Будет написан итоговый отчет. За весь период реализации проекта будут опубликованы 2 статьи в рецензируемых зарубежных научных изданиях (Scopus), в одном или нескольких из следующего списка: «Applied Artificial Intelligence», «Statistics, Optimization and Information Computing», «Journal of Artificial Intelligence», «Information and Control», «International Journal of Computing Science and Mathematics», «Applied Mathematics and Information Sciences» и 2 статьи в отечественных научных журналах с ненулевым импакт-фактором.

От Заказчика:

Председатель ГУ «Комитета науки Министерства образования и науки РК»



Абдрасилов Б.С.

М.П.

От Исполнителя:

Проректор по науке ИАО «КазНУТУ имени К.И. Сатпаева»



Кенжалиев Б.К.

М.П.

Ознакомлен:

Научный руководитель проекта(ов)

Нурсеитов Д.Б.

(подпись)