

Профессиональный стандарт: «Разработка систем обработки и хранения больших данных»

Глава 1. Общие положения

1. Область применения профессионального стандарта: Большие данные нужны в маркетинге, перевозках, автомобилестроении, здравоохранении, науке, сельском хозяйстве и других сферах, в которых можно собрать и обработать нужные массивы информации

2. В настоящем профессиональном стандарте применяются следующие термины и определения:

1) Информационная система (ИС) – организационно-упорядоченная совокупность информационно-коммуникационных технологий, обслуживающего персонала и технической документации, реализующих определенные технологические действия посредством информационного взаимодействия и предназначенных для решения конкретных функциональных задач.

2) Информационная технология (ИТ, IT) – это процесс, использующий совокупность средств и методов сбора, обработки и передачи данных для получения информации нового качества о состоянии объекта, процесса или явления. Информационные технологии (ИТ, от англ. Information Technology, IT) - это класс областей деятельности, относящихся к технологиям управления и обработкой огромного потока информации с применением вычислительной техники.

3) Сопровождение ИС – обеспечение использования введенной в промышленную эксплуатацию ИС в соответствии с ее назначением, включающее мероприятия по проведению корректировки, модификации и устранению дефектов программного обеспечения, без проведения модернизации и реализации дополнительных функциональных требований и при условии сохранения ее целостности.

4) База данных (БД) – совокупность данных, организованных согласно концептуальной структуре, описывающей характеристики этих данных, а также взаимосвязей между их объектами.

5) Глубокое обучение (англ. Deep learning) – совокупность методов машинного обучения (с учителем, с частичным привлечением учителя, без учителя, с подкреплением), основанных на обучении представлениям (англ. feature/representation learning), а не специализированным алгоритмам под конкретные задачи.

6) Искусственные нейронные сети (нейронные сети) – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма.

7) Машинное обучение (Machine Learning) – обширный подраздел искусственного интеллекта, изучающий методы построения алгоритмов, способных обучаться.

8) Хранилище данных (ХД) – предметно-ориентированная информационная база данных, специально разработанная и предназначенная для подготовки отчетов и бизнес-анализа с целью поддержки принятия решений в организации.

9) Компьютерная лингвистика (математическая или вычислительная лингвистика, англ. computational linguistics) – научное направление в области математического и компьютерного моделирования интеллектуальных процессов у человека и животных при создании систем искусственного интеллекта, которое ставит своей целью использование математических моделей для описания естественных языков.

10) Компьютерное зрение – технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов.

3. В настоящем профессиональном стандарте применяются следующие сокращения:

- 1) NLP (Natural Language Processing) – обработка естественного языка
- 2) ИКТ – Информационно-коммуникационные технологии
- 3) СУБД – Система управления базами данных

Глава 2. Паспорт профессионального стандарта

4. Название профессионального стандарта: Разработка систем обработки и хранения больших данных

5. Код профессионального стандарта: J62011034

6. Указание секции, раздела, группы, класса и подкласса согласно ОКЭД:

J Информация и связь

62 Компьютерное программирование, консультационные и другие сопутствующие услуги

62.0 Компьютерное программирование, консультационные и другие сопутствующие услуги

62.01 Деятельность в области компьютерного программирования

62.01.1 Разработка программного обеспечения

7. Краткое описание профессионального стандарта: В данном профессиональном стандарте приведены описания функций специалистов по управлению данными и специалисты NLP. Специалисты данных находят и интерпретируют богатые источники данных, управляют большими объемами данных, объединяют источники данных, обеспечивают согласованность наборов данных и создают визуализации, помогающие в понимании

данных. Они строят математические модели, используя данные, представляют их. Специалисты NPL работают в области компьютерных наук, а точнее в области обработки естественного языка. Они нацелены на то, чтобы сократить разрыв в переводе между точными переводами, выполняемыми человеком, для машинных переводчиков. Они анализируют тексты, сравнивают и отображают переводы и улучшают лингвистику переводов с помощью программирования и кода.

8. Перечень карточек профессий:

- 1) Инженер NLP (специалист по компьютерной лингвистике) - 6 уровень ОРК
- 2) Специалист по Data Mining - 6 уровень ОРК
- 3) Программист компьютерного зрения - 6 уровень ОРК
- 4) Специалист по нейронным сетям - 6 уровень ОРК
- 5) Специалист по машинному обучению - 6 уровень ОРК
- 6) Специалист по нейронным сетям - 7 уровень ОРК
- 7) Специалист по машинному обучению - 7 уровень ОРК
- 8) Инженер NLP (специалист по компьютерной лингвистике) - 7 уровень ОРК
- 9) Программист компьютерного зрения - 7 уровень ОРК
- 10) Специалист по Data Mining - 7 уровень ОРК

Глава 3. Карточки профессий

9. Карточка профессии «Инженер NLP (специалист по компьютерной лингвистике)»:			
Код группы:	2643-9		
Код наименования занятия:	2643-9-003		
Наименование профессии:	Инженер NLP (специалист по компьютерной лингвистике)		
Уровень квалификации по ОРК:	6		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: высшее образование (бакалавриат, специалитет, ординатура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Разработка алгоритмов обработки текстовой информации		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Обработка текстовой информации посредством вычислительных средств и технологии	
	Дополнительные трудовые функции:		
Трудовая функция 1: Обработка текстовой информации посредством вычислительных средств и технологии	Навык 1: Подготовка словарей для обработки данных	Умения:	
		1. Проводить выборку требуемых слов/основ для составления словаря – базы данных 2. Создавать базу данных слова для обработки языка, проверять слова на правильность использования 3. Применять орфографические, синтаксические, грамматические правила естественного языка.	

		Знания:
		<ol style="list-style-type: none"> 1. Языки программирования R, Java, C#/C++, Scala и другие 2. Математика, теория вероятностей, статистика, знание сфер применимости, понимание плюсов и минусов различных семейств алгоритмов машинного обучения (таких как логистическая регрессия, различные алгоритмы кластеризации, нейронные сети, бустинг)
	Возможность признания навыка:	-
	Навык 2: Разработка оптимальных алгоритмов для работы с текстовыми данным	Умения:
		<ol style="list-style-type: none"> 1. Разрабатывать программу для составления словаря 2. Определять пригодность данных для применения в системах обработки текстовых данных естественного языка 3. Использовать алгоритмы для автоматической обработки текста
		Знания:
		<ol style="list-style-type: none"> 1. Методы семантического, синтаксического, морфологического анализа текстовых данных 2. Методы автоматической обработки данных естественного языка
	Возможность признания навыка:	-
	Навык 3: Разработка программ для обработки текстовой информации	Умения:
		<ol style="list-style-type: none"> 1. Использовать языки программирования для разработки приложений для обработки текстовой информации 2. Определять алгоритмы обработки данных естественного языка (словарей-переводчиков, поисковиков и т.д) 3. Проводить контекстно-зависимый анализ текста. 4. Применять объекты чтения корпуса HTML, из базы данных.
		Знания:
		<ol style="list-style-type: none"> 1. Языки программирования 2. Алгоритмы для обработки текстовой информации. 3. Методы контекстно-зависимого анализа
	Возможность признания навыка:	-
Требования к личностным компетенциям:	<p> Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в принятии решения </p>	
Список технических регламентов и национальных стандартов:		
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:
	6	Специалист по DataMining
	7	Специалист по DataMining
	6	Специалист по нейронным сетям
	7	Специалист по нейронным сетям
6	Специалист по машинному обучению	

	7	Специалист по машинному обучению	
	6	Программист компьютерного зрения	
	7	Программист компьютерного зрения	
10. Карточка профессии «Специалист по Data Mining»:			
Код группы:	2521-1		
Код наименования занятия:	2521-1-003		
Наименование профессии:	Специалист по Data Mining		
Уровень квалификации по ОРК:	6		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: высшее образование (бакалавриат, специалитет, ординатура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Создание программных средств анализа больших данных и прогнозирования бизнес-процессов		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Проведение анализа огромных массивов информации 2. Разработка и управление программными средствами автоматизации обработки больших данных	
	Дополнительные трудовые функции:		
Трудовая функция 1: Проведение анализа огромных массивов информации	Навык 1: Организация обработки больших данных	Умения:	
		1. Определять критерий данных для поиска и извлечения данных 2. Применять способы разработки SQL запросов 3. Выявлять тривиальные и нетривиальные зависимости между данными 4. Генерировать релевантную информацию из извлеченных данных 5. Сегментировать группы потребителей с учетом необходимых характеристик на основе анализа полученных данных	
	Возможность признания навыка:	Знания:	
1. Математика, дискретный анализ, математическая статистику 2. Методика машинного обучения и визуализации, а также деревья решений, генетические алгоритмы, нейронные сети, ассоциативные связи, кластерный анализ 3. Способы создания запросов на языке SQL			
Трудовая функция 2: Разработка и управление программными средствами			

автоматизации обработки больших данных	Навык 1: Обеспечение хранения больших данных	Умения: 1. Создавать SQL запросов к большим данным по запрашиваемой информации с применением операторов AND и OR. 2. Создавать базы данных и хранилище данных, использовать функциональные возможности СУБД 3. Устанавливать, производить настройку конфигурации программных средств
		Знания: 1. Современные программные приложения для обработки данных 2. Методы поиска в современных базах данных 3. Методология анализа больших данных
	Возможность признания навыка:	-
	Навык 2: Создание программных средств обработки данных	Умения: 1. Определять типы и структуру обрабатываемых данных 2. Разрабатывать алгоритмы для обработки больших данных 3. Разработать и использовать методы быстрого поиска и обработки данных 4. Разрабатывать модели и способы решения задачи анализа данных
		Знания: 1. Языки программирования для анализа данных 2. Основы алгоритмизации и структуры данных, основы программирования, алгоритмы поиска больших данных
	Возможность признания навыка:	-
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в решении проблем	
Список технических регламентов и национальных стандартов:		
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:
	6-7	Специалист по нейронным сетям
	6-7	Специалист по машинному обучению
	6-7	Инженер NLP (специалист по компьютерной лингвистике)
	6-7	Программист компьютерного зрения
11. Карточка профессии «Программист компьютерного зрения»:		
Код группы:	2519-9	
Код наименования занятия:	2519-9-002	
Наименование профессии:	Программист компьютерного зрения	
Уровень квалификации по ОРК:	6	
подуровень квалификации по ОРК:		

Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: высшее образование (бакалавриат, специалитет, ординатура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Создание программных средств для компьютерного зрения		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Подготовка данных и разработка программ для обработки видео и графических изображений 2. Управление приложением и оборудованием для компьютерного зрения	
	Дополнительные трудовые функции:		
Трудовая функция 1: Подготовка данных и разработка программ для обработки видео и графических изображений	Навык 1: Выборка графических данных	Умения:	
		1. Проводить анализ видео и изображения 2. Применять готовую библиотеку TensorFlow для анализа изображения 3. Проводить кластеризацию графических данных. 4. Определять методы и алгоритмы для обнаружения и отслеживания признаков. 5. Выбирать оптимальные алгоритмы обнаружения объектов для ПО	
		Знания:	
		1. Теорию и технологию создания искусственных систем 2. Виды компьютерной графики, программные средства для создания графических объектов 3. Алгоритмы прослеживания признаков, виды признаков. 4. Методы обнаружения объектов	
	Возможность признания навыка:	-	
Трудовая функция 2:	Навык 2: Разработка программ для компьютерного зрения	Умения:	
		1. Создавать искусственные системы для обработки графических и видео данных 2. Проводить анализ графических и видео данных 3. Проводить сегментацию изображения 4. Применять метод динамического программирования для сопоставления стереоизображения	
		Знания:	
	1. Актуальные программные средства и оборудования в области компьютерного зрения и требования к ним 2. Методы и алгоритмы сегментации изображения 3. Метод динамического программирования для сопоставления стереоизображения		
Возможность признания навыка:	-		

Управление приложением и оборудованием для компьютерного зрения	Навык 1: Сопровождение ПО и устройств для компьютерного зрения	Умения:	
		1. Настраивать, конфигурировать ПО для машинного зрения 2. Производить обновление и пополнение данных 3. Производить инсталляцию программы 4. Производить установку устройств для машинного зрения	
	Возможность признания навыка:	Знания:	
1. Датчики для компьютерного зрения 2. Программное обеспечения для компьютерного зрения			
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Организованность Исполнительность Математическое мышление Гибкость мышления Ориентированность на результат Креативность Самостоятельность в принятии решения		
Список технических регламентов и национальных стандартов:			
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:	
	6	Специалист по DataMinin	
	7	Специалист по DataMinin	
	6	Специалист по нейронным сетям	
	7	Специалист по нейронным сетям	
	6	Специалист по машинному обучению	
	7	Специалист по машинному обучению	
	6	Инженер NLP (специалист по компьютерной лингвистике)	
7	Инженер NLP (специалист по компьютерной лингвистике)		
12. Карточка профессии «Специалист по нейронным сетям»:			
Код группы:	3122-2		
Код наименования занятия:	3122-2-008		
Наименование профессии:	Специалист по нейронным сетям		
Уровень квалификации по ОРК:	6		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: высшее образование (бакалавриат, специалитет, ординатура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			

Другие возможные наименования профессии:		
Основная цель деятельности:	Разработка алгоритмов решения сложных задач на основе нейросетей	
Описание трудовых функций		
Перечень трудовых функций:	Обязательные трудовые функции:	1. Подготовка данных для применения в нейросистемах 2. Применение нейронных сетей в решении сложных задач при обработке данных
	Дополнительные трудовые функции:	
Трудовая функция 1: Подготовка данных для применения в нейросистемах	Навык 1: Создание эталонной базы данных	Умения: 1. Производить сбор и систематизацию данных для применения в БД 2. Проводить моделирование перехода данных с состояния пригодности в состояние стоимости.
		Знания: 1. Теоретические основы нейронных сетей 2. Алгоритмы реализации нейронных сетей
	Возможность признания навыка:	-
	Навык 2: Моделирование систем с применением нейросетей	Умения: 1. Выбирать способы разработки модели с распределением данных по параметрам. 2. Разработать математическую модель системы 3. Определить запросы данных для их обработки на основе нейронных сетей.
		Знания: 1. Современные программные приложения 2. Языки программирования для нейросетей и анализа данных 3. Архитектура нейронных сетей
	Возможность признания навыка:	-
Трудовая функция 2: Применение нейронных сетей в решении сложных задач при обработке данных	Навык 1: Разработка программ на основе нейронных сетей	Умения: 1. Проектировать и анализировать предметную область 2. Определять алгоритмы для построения нейронных систем относительно цели и задачи работы 3. Создавать модель для реализации программы 4. Проводить верификацию код программы на правильность выполнения 5. Писать код программы для обработки данных 6. Реализовать методы и технологии глубокого обучения в разрабатываемой программе
		Знания: 1. Алгоритмы и структуры данных, основы программирования 2. Теоретические основы нейронных сетей и машинного обучения
	Возможность признания навыка:	-

Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в принятии решения		
Список технических регламентов и национальных стандартов:			
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:	
	6	Специалист по DataMining	
	7	Специалист по DataMining	
	6	Специалист по машинному обучению	
	7	Специалист по машинному обучению	
	6	Инженер NLP (специалист по компьютерной лингвистике)	
	7	Инженер NLP (специалист по компьютерной лингвистике)	
	6	Программист компьютерного зрения	
7	Программист компьютерного зрения		
13. Карточка профессии «Специалист по машинному обучению»:			
Код группы:	2422-2		
Код наименования занятия:	2422-2-009		
Наименование профессии:	Специалист по машинному обучению		
Уровень квалификации по ОРК:	6		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: высшее образование (бакалавриат, специалитет, ординатура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Разработка алгоритмов для реализации методов машинного обучения		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Проектирование и реализация систем с применением машинного обучения	
	Дополнительные трудовые функции:		
Трудовая функция 1: Проектирование и реализация систем с применением машинного обучения			

	Навык 1: Построение модели систем машинного обучения	Умения: 1. Определять задачи реализации программ машинного обучения 2. Понимать алгоритмы для решения задач по машинному обучению 3. Определять типы данных 4. Установить причинно-следственную связь между данными
		Знания: 1. Алгоритмы, применяемые в машинном обучении 2. Методы и виды машинного обучения
	Возможность признания навыка:	-
	Навык 2: Применение разработанных технологии для машинного обучения и разработка программных средств для анализа данных	Умения: 1. Применять методы машинного обучения при использовании систем для решения задач по анализу данных 2. Определить данные, классифицировать данные, определять взаимосвязь между данными для решения задач по их обработке.
		Знания: 1. Современные приложения для проведения классификации данных 2. Методы и виды машинного обучения
	Возможность признания навыка:	-
	Навык 3: Построение модели стратегической оценки, управление данными в искусственном интеллекте	Умения: 1. Определять алгоритмы для классификации данных 2. Разрабатывать программу принятия решения 3. Применять виды машинного обучения при анализе данных 4. Производить множество расчетов и сравнений с данными 5. Применять обучение по прецедентам, или индуктивное обучение, основанное на выявлении эмпирических закономерностей в данных и дедуктивное обучение. 6. Переносить базу знаний экспертов в базу данных.
		Знания: 1. Алгоритмы обработки больших данных 2. Кластеризация и классификация данных 3. Методы искусственного интеллекта 4. Математическая статистика, численные методы, методы оптимизации, теории вероятностей, теории графов, техника работы с данными в цифровой форме.
	Возможность признания навыка:	-
	Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в принятии решения
Список технических регламентов и национальных стандартов:		
Связь с другими	Уровень ОРК:	Наименование профессии:

профессиями в рамках ОРК:	6	Специалист по DataMining
	7	Специалист по DataMining
	6	Специалист по нейронным сетям
	7	Специалист по нейронным сетям
	6	Инженер NLP (специалист по компьютерной лингвистике)
	7	Инженер NLP (специалист по компьютерной лингвистике)
	6	Программист компьютерного зрения
	7	Программист компьютерного зрения

14. Карточка профессии «Специалист по нейронным сетям»:

Код группы:	3122-2		
Код наименования занятия:	3122-2-008		
Наименование профессии:	Специалист по нейронным сетям		
Уровень квалификации по ОРК:	7		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: послевузовское образование (магистратура, резидентура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Разработка алгоритмов решения сложных задач на основе нейросетей		

Описание трудовых функций

Перечень трудовых функций:	Обязательные трудовые функции:	1. Подготовка данных для применения в нейросистемах 2. Применение нейронных сетей в решении сложных задач при обработке данных
	Дополнительные трудовые функции:	
Трудовая функция 1: Подготовка данных для применения в нейросистемах	Навык 1: Создание эталонной базы данных	Умения: 1. Исследовать пути решения задачи с помощью нейронной сети 2. Скачивать, собирать, выбирать данные для работы 3. Проводить кластерный анализ данных на выходных параметрах нейронной сети
		Знания: 1. Математическая статистика, дискретная математика, математическая логика 2. Основы алгоритмизации, структуры данных и программирование.
	Возможность признания навыка:	-

	Навык 2: Моделирование систем с применением нейросетей	Умения: 1. Применять вычислительные процессы нейронной сети 2. Описать данные и их взаимодействие в системе для разработки его модели 3. Определять применяемый тип модели для реализуемой системы. 4. Исследовать виды моделей представления данных 5. Применять метод глубокого обучения
		Знания: 1. Современные программные приложения 2. Методы глубокого обучения в нейронной сети 3. Основные понятия моделирования и виды моделей
	Возможность признания навыка:	-
Трудовая функция 2: Применение нейронных сетей в решении сложных задач при обработке данных	Навык 1: Разработка программ на основе нейронных сетей	Умения: 1. Разрабатывать проект реализации программы для обработки графической и визуальной информации 2. Определять математическую модель программы 3. Разрабатывать алгоритм и методы реализации вычислительных систем с нейронными сетями. 4. Использовать готовые алгоритмы нейронных сетей (программные средства), библиотеки для обработки данных 5. Выбрать язык программирования для работы с нейронными сетями 6. Использовать средства с технологией нейронных сетей для обработки текстовых, звуковых, графических данных
		Знания: 1. Принципы параллельной обработки данных информационных систем 2. Нейросетевое распознавание символов текста многослойными перцептронами. 3. Современные средства с реализацией технологией нейронных сетей 4. Архитектура нейронных сетей
	Возможность признания навыка:	-
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в принятии решения	
Список технических регламентов и национальных стандартов:		
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:
	6	Специалист по DataMining
	7	Специалист по DataMining
	6	Специалист по машинному обучению
	7	Специалист по машинному обучению
	6	Инженер NLP (специалист по компьютерной лингвистике)
	7	Инженер NLP (специалист по компьютерной лингвистике)
	6	Программист компьютерного зрения
7	Программист компьютерного зрения	

15. Карточка профессии «Специалист по машинному обучению»:

Код группы:	2422-2		
Код наименования занятия:	2422-2-009		
Наименование профессии:	Специалист по машинному обучению		
Уровень квалификации по ОРК:	7		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: послевузовское образование (магистратура, резидентура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Разработка алгоритмов для реализации методов машинного обучения		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Проектирование и реализация систем с применением машинного обучения	
	Дополнительные трудовые функции:		
Трудовая функция 1: Проектирование и реализация систем с применением машинного обучения	Навык 1: Построение модели систем машинного обучения	Умения:	
		1. Определять алгоритмы для создания модели: наивная байесовская классификация, дерево принятия решений, метод наименьших квадратов, метод опорных векторов (SVM) и другие 2. Подготовить обработанные данные для создания модели обучения 3. Проводить оценку качества данных и подготовить аналитический отчет по результатам работы.	
		Знания:	
	1. Алгоритмы, применяемые в машинном обучении. 2. Современные системы с реализованной технологией машинного обучения. 3. Дискретная математика, статистический анализ данных, теория вероятностей.		
	Возможность признания навыка:	-	
	Навык 2: Применение разработанных технологии для машинного обучения и разработка программных средств для анализа данных	Умения:	
		1. Производить обработку и подготовку данных для машинного обучения. 2. Выбирать алгоритмы для обработки данных путем машинного обучения. 3. Выбирать язык программирования для реализации программы.	

		Знания: 1. Алгоритмы машинного обучения для обработки данных 2. Языки программирования для анализа данных и обработки данных 3. Математическая логика, математическая статистика и дискретная математика
	Возможность признания навыка:	-
	Навык 3: Построение модели стратегической оценки, управление данными в искусственном интеллекте	Умения: 1. Определять алгоритмы для классификации данных 2. Разрабатывать программу принятия решения 3. Применять виды машинного обучения при анализе данных 4. Производить множество расчетов и сравнений с данными 5. Применять обучение по прецедентам, или индуктивное обучение, основанное на выявлении эмпирических закономерностей в данных и дедуктивное обучение. 6. Переносить базу знаний экспертов в базу данных. Знания: 1. Алгоритмы обработки больших данных 2. Кластеризация и классификация данных 3. Методы искусственного интеллекта 4. Математическая статистика, численные методы, методы оптимизации, теории вероятностей, теории графов, техника работы с данными в цифровой форме.
	Возможность признания навыка:	-
Требования к личностным компетенциям:	<p>Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в решении проблем</p>	
Список технических регламентов и национальных стандартов:		
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:
	6	Специалист по DataMining
	7	Специалист по DataMining
	6	Специалист по нейронным сетям
	7	Специалист по нейронным сетям
	6	Инженер NLP (специалист по компьютерной лингвистике)
	7	Инженер NLP (специалист по компьютерной лингвистике)
	6	Программист компьютерного зрения
7	Программист компьютерного зрения	
16. Карточка профессии «Инженер NLP (специалист по компьютерной лингвистике)»:		
Код группы:	2643-9	
Код наименования занятия:	2643-9-003	
Наименование профессии:	Инженер NLP (специалист по компьютерной лингвистике)	
Уровень квалификации по ОРК:	7	
подуровень квалификации по ОРК:		

Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: послевузовское образование (магистратура, резидентура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Разработка алгоритмов обработки текстовой информации		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Обработка текстовой информации посредством вычислительных средств и технологии	
	Дополнительные трудовые функции:		
Трудовая функция 1: Обработка текстовой информации посредством вычислительных средств и технологии	Навык 1: Подготовка словарей для обработки данных	Умения:	
			1. Проводить аналитические процессы с текстовыми данными 2. Определять пригодность текстовых данных, 3. Применять методы обработки текстовых данных 4. Применять алгоритмы компьютерной лингвистики при обработке текста
		Знания:	
			1. Проводить анализ текстовых данных 2. Определять пригодность текстовых данных 3. Применять методы обработки текстовых данных 4. Изучать алгоритмы, применяемые в компьютерной лингвистике
	Возможность признания навыка:	-	
	Навык 2: Разработка оптимальных алгоритмов для работы с текстовыми данным	Умения:	
		1. Исследовать методы разработки алгоритмов для быстрой обработки естественного языка 2. Разрабатывать приложения для проведения семантического, синтаксического, морфологического анализа естественного языка 3. Сравнивать и анализировать эффективность применения разработанных алгоритмов	
Знания:			
		1. Технология разработки современных приложений для обработки естественного языка 2. Современные языки программирования. 3. Алгоритмы для обработки естественного языка	
Возможность признания навыка:	-		

	Навык 3: Разработка программ для обработки текстовой информации	Умения: 1. Применять для решения задач по обработке текста глубокое обучение, нейронные сети. 2. Применять конвейеры векторизации и преобразования корпуса текста 3. Производить кластеризацию данных для выявления сходства в тексте.	
	Возможность признания навыка:	Знания: 1. Алгоритм Портера для составления словаря из основ слов 2. Нейронные сети, глубокое обучение, машинное обучение 3. Виды кластеризации данных	
		-	
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Ориентированность на результат Организованность Креативность самостоятельность в принятии решения		
Список технических регламентов и национальных стандартов:			
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:	
	6	Специалист по DataMining	
	7	Специалист по DataMining	
	6	Специалист по нейронным сетям	
	7	Специалист по нейронным сетям	
	6	Специалист по машинному обучению	
	7	Специалист по машинному обучению	
	6	Программист компьютерного зрения	
7	Программист компьютерного зрения		
17. Карточка профессии «Программист компьютерного зрения»:			
Код группы:	2519-9		
Код наименования занятия:	2519-9-002		
Наименование профессии:	Программист компьютерного зрения		
Уровень квалификации по ОРК:	7		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: послевузовское образование (магистратура, резидентура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			

Другие возможные наименования профессии:		
Основная цель деятельности:	Создание программных средств для компьютерного зрения	
Описание трудовых функций		
Перечень трудовых функций:	Обязательные трудовые функции:	<ol style="list-style-type: none"> 1. Подготовка данных и разработка программ для обработки видео и графических изображений 2. Управление приложением и оборудованием для компьютерного зрения
	Дополнительные трудовые функции:	
Трудовая функция 1: Подготовка данных и разработка программ для обработки видео и графических изображений	Навык 1: Выборка графических данных	Умения:
		<ol style="list-style-type: none"> 1. Анализировать область применения системы для компьютерного зрения 2. Проводить исследования и разработать алгоритмы в областях: технического зрения, 3D-реконструкции по изображениям, анализа сцен и распознавания, обработки изображений 3. Применять анализ изображений и алгоритмы анализа плотного движения.
		Знания:
		<ol style="list-style-type: none"> 1. Устройства для компьютерного зрения 2. Программные обеспечения для компьютерного зрения 3. Алгоритмы для разработки приложений для компьютерного зрения 4. Методы анализ изображений и алгоритмы анализа плотного движения (Алгоритмы Хорна - Шанка, Лукаса – Канаде, BBPW, вычисления оптического потока).
	Возможность признания навыка:	-
	Навык 2: Разработка программ для компьютерного зрения	Умения:
<ol style="list-style-type: none"> 1. Проектировать архитектуру системы компьютерного зрения 2. Создавать прототип ПО для компьютерного зрения 3. Производить трехмерную реконструкцию изображений 		
Знания:		
	<ol style="list-style-type: none"> 1. Современные программные приложения 2. Средства моделирования ПО 3. Методы и принципы проектирования ПО 4. Основы стереоскопического зрения, фотометрический метод анализа стереоизображений 	
Возможность признания навыка:	-	
Трудовая функция 2: Управление приложением и оборудованием для компьютерного зрения	Навык 1: Сопровождение ПО и устройств для компьютерного зрения	Умения:
		<ol style="list-style-type: none"> 1. Определять место установки камер видеорегистрации данных 2. Обеспечить работоспособность программы и системы компьютерного зрения, проводить мониторинг системы. 3. Обнаруживать ошибки и вносить изменения, производить замену ПО для компьютерного зрения
		Знания:
		<ol style="list-style-type: none"> 1. Основы работ IoT систем 2. Устройства для компьютерного зрения 3. Методика и принципы работы устройств компьютерного зрения

	Возможность признания навыка:	-	
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Гибкость мышления Ориентированность на результат Организованность Креативность Самостоятельность в принятии решения		
Список технических регламентов и национальных стандартов:			
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:	
	6	Специалист по Data Mining	
	7	Специалист по Data Mining	
	6	Специалист по нейронным сетям	
	7	Специалист по нейронным сетям	
	6	Специалист по машинному обучению	
	7	Специалист по машинному обучению	
	6	Инженер NLP (специалист по компьютерной лингвистике)	
7	Инженер NLP (специалист по компьютерной лингвистике)		
18. Карточка профессии «Специалист по Data Mining»:			
Код группы:	2521-1		
Код наименования занятия:	-		
Наименование профессии:	Специалист по Data Mining		
Уровень квалификации по ОРК:	7		
подуровень квалификации по ОРК:			
Уровень квалификации по ЕТКС, КС и др типовых квалификационных характеристик:	Квалификационный справочник должностей руководителей, специалистов и иных служащих Приказ Министра труда и социальной защиты населения Республики Казахстан от 30 декабря 2020 года № 553 "Об утверждении Квалификационного справочника должностей руководителей, специалистов и других служащих". Зарегистрирован в Министерстве юстиции Республики Казахстан 31 декабря 2020 года № 22003. 140. Инженер-программист 96. Руководитель проекта		
Уровень профессионального образования:	Уровень образования: послевузовское образование (магистратура, резидентура)	Специальность: Информационно-коммуникационные технологии	Квалификация: -
Требования к опыту работы:			
Связь с неформальным и информальным образованием:			
Другие возможные наименования профессии:			
Основная цель деятельности:	Создание программных средств анализа больших данных и прогнозирования бизнес-процессов		
Описание трудовых функций			
Перечень трудовых функций:	Обязательные трудовые функции:	1. Проведение анализа огромных массивов информации. 2. Разработка и управление программными средствами автоматизации обработки больших данных	
	Дополнительные трудовые функции:		
Трудовая функция 1: Проведение анализа огромных массивов информации.			

	<p>Навык 1: Организация обработки больших данных</p>	<p>Умения:</p> <ol style="list-style-type: none"> 1. Определять способы и методы работы по извлечению ранее неизвестных данных для получения знаний 2. Построить SQL запросы для извлечения знаний. 4. Применять при обработке данных методы классификации, кластеризации, регрессии, ассоциации, анализ отклонений, последовательного шаблона. 5. Разработать методы по проведению анализа данных, которые используются для разработки корпоративных систем и приложений для получения знаний и информации, улучшающих бизнес-процессы. 6. Производить математические расчеты для извлечения знаний. <p>Знания:</p> <ol style="list-style-type: none"> 1. Математическая статистика, математическую логику 2. Средства работы с SQL и язык SQL 3. Методологию аналитических исследований и методы прогнозирования в маркетинге (сбыт, конкурентноспособность и так далее).
	<p>Возможность признания навыка:</p>	<p>-</p>
<p>Трудовая функция 2: Разработка и управление программными средствами автоматизации обработки больших данных</p>	<p>Навык 1: Обеспечение хранения больших данных</p>	<p>Умения:</p> <ol style="list-style-type: none"> 1. Проводить оценку потребностей 2. Определить этапы принятия решений в ключевых бизнес-процессах 3. Определить бизнес-процессы для проведения анализа 4. Собрать данные и провести анализ данных при помощи интеллектуальных систем 1. Определять способы работы и методы по извлечению ранее неизвестных данных для получения знаний 2. Применять SQL запросы для извлечения знаний 3. Проводить кластерный анализ больших данных 5. Разработать методы по проведению анализа данных, которые используются для разработки корпоративных систем и приложений для получения знаний и информации, улучшающих бизнес-процессы <p>Знания:</p> <ol style="list-style-type: none"> 1. Математическая статистика, математическая логика, машинное обучение 2. Методы и виды прогнозирования 3. Интеллектуальных систем анализа данных
	<p>Возможность признания навыка:</p>	<p>-</p>
	<p>Навык 2: Создание программных средств обработки данных</p>	<p>Умения:</p> <ol style="list-style-type: none"> 1. Решать задачи по анализу больших данных при помощи языка программирования. 2. Использовать метаданные для работы при создании файлов или таблиц базы данных 3. Применять методы искусственного интеллекта в программировании. <p>Знания:</p> <ol style="list-style-type: none"> 1. Современные приложения для проведения интеллектуального анализа данных, 2. Современные методы и технологии искусственного интеллекта.

	Возможность признания навыка:	-
Требования к личностным компетенциям:	Ответственность Аналитическое мышление Логическое мышление Исполнительность Математическое мышление Ориентированность на результат Организованность Креативность Самостоятельность в решении проблем	
Список технических регламентов и национальных стандартов:		
Связь с другими профессиями в рамках ОРК:	Уровень ОРК:	Наименование профессии:
	6-7	Специалист по нейронным сетям
	6-7	Специалист по машинному обучению
	6-7	Инженер NLP (специалист по компьютерной лингвистике)
	6-7	Программист компьютерного зрения

Глава 4. Технические данные профессионального стандарта

19. Наименование государственного органа:

20. Организации (предприятия) участвующие в разработке:

Товарищество с ограниченной ответственностью «Компания системных исследований «Фактор»

Руководитель проекта:

Габбасов М.Б.

E-mail: Mars0@mail.ru

Номер телефона: +7 (701) 908 25 11

Исполнители:

Исин Н.К., +7 (701) 111 18 71, info@itk.kz

Абдешов Х.У., +7 (777) 250 58 31, habdeshov@rambler.ru

Аканова А.С., +7 (705) 448 06 80, akerkegansaj@mail.ru

21. Отраслевой совет по профессиональным квалификациям:

22. Национальный орган по профессиональным квалификациям: -

23. Национальная палата предпринимателей Республики Казахстан «Атамекен»: 24.12.2019 г.

24. Номер версии и год выпуска: версия 2, 2022 г.

25. Дата ориентировочного пересмотра: 30.12.2025 г.