Изображение выглядит как текст, Шрифт, логотип, Графика

Автоматически созданное описание

КОНКУРСНОЕ ЗАДАНИЕ КОМПЕТЕНЦИИ

«НЕЙРОСЕТИ И БОЛЬШИЕ ДАННЫЕ»

г. Великий Новгород

2023г.

Конкурсное задание разработано экспертным сообществом и утверждено Менеджером компетенции, в котором установлены нижеследующие правила и необходимые требования владения профессиональными навыками для участия в соревнованиях по профессиональному мастерству.

**Конкурсное задание включает в себя следующие разделы:**

[1. Основные требования компетенции 4](#_Toc132027311)

[1.1. Общие сведения о требованиях компетенции 4](#_Toc132027312)

[1.2. Перечень профессиональных задач специалиста по компетенции «Нейросети и большие данные» 4](#_Toc132027313)

[1.3. Требования к схеме оценки 6](#_Toc132027314)

[1.4. Спецификация оценки компетенции 6](#_Toc132027315)

[2. Специальные правила компетенции 11](#_Toc132027317)

[2.1. Личный инструмент конкурсанта 12](#_Toc132027318)

[2.2.Оборудование запрещенное на площадке 12](#_Toc132027319)

[3. Приложения 13](#_Toc132027320)

**ИСПОЛЬЗУЕМЫЕ СОКРАЩЕНИЯ**

1. API – Application Programming Interface – интерфейс прикладного программирования;
2. IP-камера – камера, использующая IP-протокол (Internet Protocol – межсетевой протокол взаимодействия) для передачи данных;
3. USB-накопитель – запоминающее устройство, использующее в качестве носителя флеш-память, и подключаемое к компьютеру или иному считывающему устройству по интерфейсу USB (Universal Serial Bus – универсальная последовательная шина);
4. КЗ – конкурсное задание;
5. СПО – среднее профессиональное образование;
6. НСиБД – Нейросети и большие данные

1. ОСНОВНЫЕ ТРЕБОВАНИЯ КОМПЕТЕНЦИИ

1.1. ОБЩИЕ СВЕДЕНИЯ О ТРЕБОВАНИЯХ КОМПЕТЕНЦИИ

Требования компетенции «Нейросети и большие данные» определяют знания, умения, навыки и трудовые функции, которые лежат в основе наиболее актуальных требований работодателей отрасли.

Целью соревнований по компетенции является демонстрация лучших практик и высокого уровня выполнения работы по соответствующей рабочей специальности или профессии.

Требования компетенции являются руководством для подготовки конкурентоспособных, высококвалифицированных специалистов / рабочих и участия их в конкурсах профессионального мастерства.

В соревнованиях по компетенции проверка знаний, умений, навыков и трудовых функций осуществляется посредством оценки выполнения практической работы.

Требования компетенции разделены на четкие разделы с номерами и заголовками, каждому разделу назначен процент относительной важности, сумма которых составляет 100.

## 1.2. ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ ЗАДАЧ СПЕЦИАЛИСТА ПО КОМПЕТЕНЦИИ «Нейросети и большие данные»

*Таблица №1*

**Перечень профессиональных задач специалиста**

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 1 | **Разработка модулей программного обеспечения для компьютерных систем** | 30% |
| Специалист должен знать и понимать:  - основные этапы разработки программного обеспечения;  - основные принципы технологии структурного и объектно-ориентированного программирования;  - способы оптимизации и приемы рефакторинга;  - основные принципы отладки и тестирования программных продуктов. |
| Специалист должен уметь:  - осуществлять разработку кода программного модуля на языках низкого и высокого уровней;  - создавать программу по разработанному алгоритму как отдельный модуль;  - выполнять отладку и тестирование программы на уровне модуля;  - осуществлять разработку кода программного модуля на современных языках программирования;  - уметь выполнять оптимизацию и рефакторинг программного кода;  - оформлять документацию на программные средства. |

*Продолжение таблицы №1*

**Перечень профессиональных задач специалиста**

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
| 2 | **Разработка и отладка программного кода** | 30% |
| Специалист должен знать и понимать:  - методы и приемы формализации поставленных задач;  - методы и приемы алгоритмизации поставленных задач;  - синтаксис выбранного языка программирования, особенности программирования на этом языке, стандартные библиотеки языка программирования;  - методологии разработки компьютерного программного обеспечения  - методы повышения читаемости программного кода;  - основные стандарты оформления технической документации на компьютерное программное обеспечение;  - методы и приемы отладки программного кода. |
| Специалист должен уметь:  - применять алгоритмы решения типовых задач в области разработки;  - использовать методы и приемы формализации поставленных задач;  - использовать методы и приемы алгоритмизации поставленных задач;  - применять выбранные языки программирования для написания программного кода;  - использовать выбранную среду программирования;  - использовать возможности имеющейся технической и/или программной архитектуры для написания программного кода;  - применять нормативно-технические документы, определяющие требования к оформлению программного кода;  - выявлять ошибки в программном коде;  - применять методы и приемы отладки программного кода;  - интерпретировать сообщения об ошибках, предупреждения, записи технологических журналов; |
| 3 | **Анализ данных и построение моделей машинного обучения** | 40% |
| Специалист должен знать и понимать:  - приемы анализа данных;  - принципы работы с графическими данными;  - принципы работы с текстовыми данными;  - принципы работы с аудио данными;  - принципы работы с видео данными;  - различные методы и алгоритмы машинного обучения;  - критерии качества моделей машинного обучения;  - последовательность разработки моделей машинного обучения;  - какие методы машинного обучения применять в зависимости от исходных данных;  - как работать с различными выборками данных;  - как использовать различные программные средства для разработки и улучшения моделей. |

*Окончание таблицы №1*

**Перечень профессиональных задач специалиста**

|  |  |  |
| --- | --- | --- |
| **№** | **Раздел** | **Важность в %** |
|  | Специалист должен уметь:  - структурировать данные;  - проводить нормализацию и подготовку данных;  - выделять признаки, свойства и характеристики объектов в данных;  - осуществлять операции с большими данными;  - проводить визуальный анализ данных;  - применять классические алгоритмы машинного обучения:   * обучение без учителя (уменьшение размерности, поиск правил, кластеризация); * обучение с учителем (регрессия, классификация);   - применять методы глубокого обучения и Нейросети:   * Перцептроны; * Сверточные нейросети; * Реккурентные сети. |  |

## 1.3. ТРЕБОВАНИЯ К СХЕМЕ ОЦЕНКИ

Сумма баллов, присуждаемых по каждому аспекту, должна попадать в диапазон баллов, определенных для каждого раздела компетенции, обозначенных в требованиях и указанных в таблице №2.

*Таблица №2*

**Матрица пересчета требований компетенции в критерии оценки**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Критерий/Модуль** | | | | | | **Итого баллов за раздел ТРЕБОВАНИЙ КОМПЕТЕНЦИИ** |
| **Разделы ТРЕБОВАНИЙ КОМПЕТЕНЦИИ** |  | **A** | **Б** | **В** | **Г** |  |
| **1** | 10 | 10 | 10 | 0 | 30 |
| **2** | 0 | 10 | 10 | 10 | 30 |
| **3** | 20 | 20 | 0 | 0 | 40 |
| **Итого баллов за критерий/модуль** | | 30 | 40 | 20 | 10 | **100** |

1.4. СПЕЦИФИКАЦИЯ ОЦЕНКИ КОМПЕТЕНЦИИ

Оценка Конкурсного задания будет основываться на критериях, указанных в таблице №3:

*Таблица №3*

**Оценка конкурсного задания**

|  |  |  |
| --- | --- | --- |
| **Критерий** | | **Методика проверки навыков в критерии** |
| **А** | **Анализ и предобработка данных** | Экспертам необходимо сравнить анализ и подготовку данных участника с эталонным образцом, после чего происходит объективная оценка по критериям. |
| **Б** | **Разработка модели машинного обучения** | Экспертам необходимо проанализировать исходный код нейросети, обучающий набор данных и документацию, выполненные конкурсантам и произвести объективные и субъективные оценки по критериям. |
| **В** | **Тестирование разработанной модели** | Экспертам необходимо выгрузить работы участников и провести их тестирование с помощью подготовленных тест-кейсов, оценивание производится по объективным критериям. |
| **Г** | **Презентация решения** | Эксперты выслушивают презентации, оценивают содержание и выступление конкурсантов с помощью объективных и субъективных критериев. |

**1.5. КОНКУРСНОЕ ЗАДАНИЕ**

**Общая продолжительность Конкурсного задания[[1]](#footnote-1):** 18 ч.

**Количество конкурсных дней:** 3 дня

Вне зависимости от количества модулей, КЗ должно включать оценку по каждому из разделов требований компетенции.

Оценка знаний участника должна проводиться через практическое выполнение Конкурсного задания. В дополнение могут учитываться требования работодателей для проверки теоретических знаний / оценки квалификации.

**1.5.1. Разработка/выбор конкурсного задания**

Конкурсное задание состоит из 4 модулей, включает обязательную к выполнению часть (инвариант) – 3 модуля: анализ и предобработка данных, разработка модели машинного обучения, тестирование разработанной модели; и вариативная часть – 1 модуль: презентация решения. Общее количество баллов конкурсного задания составляет 100.

Обязательная к выполнению часть (инвариант) выполняется всеми регионами без исключения на всех уровнях чемпионатов.

Количество модулей из вариативной части, выбирается регионом самостоятельно в зависимости от материальных возможностей площадки соревнований и потребностей работодателей региона в соответствующих специалистах. В случае если ни один из модулей вариативной части не подходит под запрос работодателя конкретного региона, то вариативный модуль формируется регионом самостоятельно под запрос работодателя. При этом время на выполнение модуля и количество баллов в критериях оценки по аспектам не меняются.

*Таблица №4*

**Матрица конкурсного задания**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Обобщенная трудовая функция | Трудовая функция | Нормативный документ/ЗУН | Модуль | Константа/вариатив | ИЛ | КО |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Матрица конкурсного задания представлена по ссылке: <https://disk.yandex.ru/i/Ij9OEMI-siy83A>

Инструкция по заполнению матрицы конкурсного задания **(Приложение № 1)**

1.5.2. Структура модулей конкурсного задания

**Введение:** В рамках Финала Чемпионата Высоких Технологий необходимо разработать обучающее приложение, которое будет взаимодействовать с нейронными сетями от компании ПАО Сбербанк – Kandinsky 2.2 и GigaChat.

В задачи конкурсанта будет входить анализ API для работы с данными нейросетями, подготовка датасета для обработки исключительных ситуаций (например, если указанные нейронные сети будут недоступны), а также создание резервной обученной модели, которая позволит обрабатывать запросы пользователя без доступа к Kandinsky 2.2 и GigaChat.

Помимо этого предстоит разработать интерфейс, который позволит генерировать обучающие задачи с помощью нейронной сети, а также сопровождать их иллюстрациями. Необходимо учесть, что пользователь должен иметь возможность указания параметров – например, количество иллюстраций, длина текста в символах, сложность задачи и так далее.

После разработки программного обеспечения необходимо сопроводить приложение документацией и тестированием, для обеспечения качества разработанного продукта.

На финальном этапе конкурсанты должны подготовить презентацию, включающую основные этапы их работы, результаты тестирования, а также демонстрацию работы программы.

**Модуль А. Анализ и предобработка данных (инвариант)**

**Время на выполнения модуля:**6 часов

**Краткое описание задания:** в рамках данного модуля участникам необходимо подготовить набор данных для обучения дублирующей нейронной сети, а также ознакомится с предоставленными организаторами API.

**Задание:**

Так как предполагается генерация заданий с помощью сторонних нейронных сетей, необходимо предусмотреть механизмы, на случай их недоступности (например, отсутствие интернет-соединения). Поэтому важно создать простую нейросеть, которая могла бы обрабатывать запросы от пользователя в случае исключительных ситуаций – например, давать простые инструкции по устранению возможных проблем, сообщать о задержках со стороны сервера, давать советы по формированию запроса к нейронной сети и так далее. Таким образом, нейросеть-дублер должна понимать запросы пользователя и давать ответ в случае сбоев:

* в случае отсутствия интернет-соединения;
* в случае проблем с клиентом;
* в случае проблем с сервером.

Давать рекомендации по устранению вышеназванных проблем, а также:

* по формированию запросов к нейронным сетям Kandinsky 2.2 и GigaChat;
* общие рекомендации, касаемо возможностей нейронных сетей (какие предметные области доступны, какая сложность, и т.п.).

Помимо подбора данных, конкурсантам предоставляется возможность ознакомиться с API Kandinsky 2.2 и GigaChat (<https://developers.sber.ru/portal/products/gigachat?attempt=1> и <https://www.sberbank.com/promo/kandinsky>) для последующей разработки приложения на их основе.

В результате работ по данному модулю конкурсант должен подготовить набор данных для последующего обучения нейронной сети, а также описать в отчете его структуру.

*Результаты работы (описание набора данных) необходимо выгрузить в персональный репозиторий.*

**Модуль Б.** **Разработка модели машинного обучения (инвариант)**

**Время на выполнения модуля:**9 часов

**Краткое описание задания:** в рамках данного модуля необходимо разработать модель нейросети-дублера, а также разработать интерфейс приложения, взаимодействующего с нейронными сетями Kandinsky 2.2 и GigaChat.

**Задание:**

В рамках данного модуля конкурсантам необходимо решить ряд задач: первая, это разработать нейросеть-дублер и обучить ее на основе подготовленных в прошлом модуле данных.

Основной задачей нейросети-дублера является общение с пользователем в случае невозможности выполнения основной задачи приложения – генерации обучающих материалов. Основной перечень обрабатываемых задач представлен в модуле А, однако по усмотрению конкурсанта список может быть дополнен.

Второй основной задачей является разработка приложения, позволяющего с помощью API взаимодействовать с нейронными сетями Kandinsky 2.2 и GigaChat для генерации обучающего контента. Примерный алгоритм работы программы:

1. Пользователь выбирает тему, для которой требуется создать задачу.
2. Приложение проверяет наличие возможности нейронной сети генерации задачи для данной темы.
3. При выборе темы приложение может предоставить пользователю несколько опций для настройки генерации текста, таких как количество предложений, длина текста, сложность и прочее.
4. Приложение обращается к нейронной сети для генерации текста на выбранную тему, учитывая настройки, выбранные пользователем в предыдущем шаге.
5. Сгенерированный текст отображается в приложении, где пользователь может его отредактировать и дополнить, при необходимости.
6. Приложение автоматически создает визуальные элементы для обучающего материала, такие как изображения.
7. Пользователь может просмотреть и отредактировать материал перед его сохранением.
8. Приложение позволяет пользователю сохранить материал в различных форматах, таких как веб-страницы, PDF-файлы, DOCX-документы и прочие.

*Результаты работы (обученная модель, описание базы данных, исходный код интерфейса) необходимо выгрузить в персональный репозиторий.*

**Модуль В. Тестирование разработанной модели (инвариант)**

**Время на выполнения модуля:**2 часа

**Краткое описание задания:** в рамках данного модуля необходимо произвести тестирование разработанной системы.

**Задание:**

В рамках данного модуля конкурсантам необходимо протестировать работу разработанной ими нейросети-дублера, а также самого приложения. Тестирование возможно осуществить в виде тестовых сценариев и интеграционного тестирования. Основной критерий – это покрытие итогового приложения тестами и демонстрация объема выполненной работы.

*Результаты выполнения задания конкурсанты должны предоставить в виде отчетов, содержащих результаты тестирования, которые необходимо выгрузить в персональный репозиторий.*

**Модуль Г. Презентация решения (вариатив)**

**Время на выполнения модуля:**1 час

**Краткое описание задания:** в рамках данного модуля участникам необходимо подготовить презентацию своего решения, в которой необходимо отразить результаты тестирования, основные этапы работы, а также продемонстрировать работу решения в режиме реального времени.

**Задание:**

В рамках данного модуля конкурсантам необходимо подготовить презентацию реализованного ими проекта. Необходимо отразить следующие пункты: информация о данных, которые использовались для обучения нейросети-дублера; разработка нейросети-дублера; документацию к разработанному интерфейсу; результаты тестирования нейронной сети. Далее презентацию необходимо защитить перед экспертами.

*Результаты работы необходимо выгрузить в персональный репозиторий.*

## 2. СПЕЦИАЛЬНЫЕ ПРАВИЛА КОМПЕТЕНЦИИ*[[2]](#footnote-2)*

В ходе проведения чемпионата есть вероятность возникновения спорных ситуаций, дающих преимущества некоторым участникам над другими. В таблице 4 представлены наиболее частые проблемы, а также алгоритм действия, при выявлении подобных ситуаций:

*Таблица №4*

**Решение спорных ситуаций**

|  |  |
| --- | --- |
| **Вопрос** | **Специальные правила** |
| Сохранение решения конкурсантами | Решение должно быть сохранено в указанную в Конкурсном задании директорию, в случае нарушения инструкций решение не проверяется. |
| Временной регламент | 1. В случае выполнения участником модуля Конкурсного задания в непредназначенное для этого время, в том числе во время выполнения другого модуля, такое решение не будет оценено.  2. За не прекращение выполнения Конкурсного задания по команде «СТОП» Главного эксперта или ответственных за хронометраж Экспертов в соответствии с временем выполнения Модуля или в иных ситуациях к Конкурсантам применяются штрафные санкции. |
| Программное обеспечение | 1. Конкурсанты могут создавать программные продукты, оформлять инструкции или делать заметки, находясь на рабочей площадке, однако их никогда нельзя забирать с рабочей площадки.  2. За использование материалов, файлов, подготовленных вне конкурсного времени и за пределами конкурсной площадки, в том числе шпаргалок, материалов, полученных в сети Интернет (если иное не указано в Конкурсном задании), выполняемый модуль Конкурсного задания не оценивается. |
| Отказ оборудования | Если имеется явное доказательство того, что конкурсанты сами причинили ущерб оборудованию, им не будет предоставляться замена и дополнительное время. |
| Поведение конкурсантов | 1. За использование ненормативной лексики устно во время выполнения Конкурсного задания или во время защиты своих работ, а также письменно в представленных к проверке файлах к Конкурсанту применяются штрафные санкции в виде снижения общей суммы баллов на 5 баллов.  2. Общение участников во время выполнения Конкурсного задания запрещено, в случае неоднократного нарушения запрета участники дисквалифицируются на оставшееся время конкурсного дня. |

2.1. Личный инструмент конкурсанта

Конкурсант может привезти индивидуальное периферийное оборудование по списку: мышь, клавиатура, проводная гарнитура.

Ко всей периферии применяется следующее требование: оборудование не должно иметь возможности беспроводного подключения, а также заранее программируемых команд (макросов).

2.2.Оборудование запрещенное на площадке

Для участников соревнований: телефоны и иные средства связи, ноутбуки, смарт-часы, Bluetooth-гарнитуры, средства фото-, аудио-, видеозаписи, средства электронного переноса информации (USB-накопители).

Для экспертного сообщества: ограничений нет.

3. ПРИЛОЖЕНИЯ

Приложение №1 Инструкция по заполнению матрицы конкурсного задания

Приложение №2 Матрица конкурсного задания

Приложение №3 Критерии оценки

Приложение №4 Инструкция по охране труда и технике безопасности по компетенции «Нейросети и большие данные».

1. *Указывается суммарное время на выполнение всех модулей КЗ одним конкурсантом.* [↑](#footnote-ref-1)
2. *Указываются особенности компетенции, которые относятся ко всем возрастным категориям и чемпионатным линейкам без исключения.* [↑](#footnote-ref-2)