Частное техническое задание

Этап 1.5.1

Разработать ОО ПК ОГО, ПК ММР и программную документацию (ПД), включая эксплуатационную документацию по их использованию для оценки гидрометеорологической обстановки в бассейнах трансграничных рек Беларуси с использованием данных ДЗЗ и прогноза максимальных уровней воды на трансграничных реках Беларуси.

Выходная продукция:

ОО ПК ОГО.

ПД на ОО ПК ОГО.

ОО ПК ММР

ПД на ОО ПК ММР

Справка-отчет о НИОКР.

СОДЕРЖАНИЕ

**Задачи**

ПК ОГО должен представлять собой веб-приложение, состоящее из следующих компонентов:

1. База данных для ведения, хранения и обеспечения доступа к входным и выходным данным - справочной информации, результатам обучения моделей множественной регрессии, результатам вычисления прогнозов и данных выходных, данным выходных форм (отчетов);
2. Серверная часть (бэкенд):

– функционал по обучению моделей множественной регрессии;

– функционал по вычислению прогнозов на обученных моделях множественной регрессии;

– программный интерфейс приложения API, обеспечивающий взаимодействие между базой данных, функционалом обучения моделей МР и прогнозирования, пользовательским интерфейсом.

1. Пользовательский интерфейс (фронтенд):

- отображение, ввод, корректировка справочной информации;

- ввод и отображение данных, необходимых для обучения моделей множественной регрессии;

- запуск процесса обучения и отображение результатов обучения моделей множественной регрессии;

- запуск процесса прогнозирования и отображение результатов прогноза максимальных уровней воды в реках по гидропунктам

**База данных**

База данных:

postgresql

192.168.29.134:5432

Имя БД: pkogo

Схема: maxlevel

User: pkogouser

Pass: pkogouser

Таблицы.

Справочная информация:

- реки

- гидропункты

- предикторы (предикторы, разделенные на группы в соответствии с методикой расчета)

- нормы (нормированные значения предикторов)

- алгоритмы (алгоритмы множественной регрессии, используемые для формирования соответствующих моделей множественной регрессии)

Таблицы данных для обучения и прогноза по моделям множественной регрессии:

- наборы для обучения (данные, по которым обучаются модели)

- модели (обученные модели множественной регрессии с параметрами оценки качества)

- наборы для прогноза (данные, по которым вычисляются прогнозы)

- прогнозы (результаты запуска обученных моделей)

- оценки качества моделей множественной регрессии

- проверочные прогнозы

Таблица реки (rivers)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| river\_id | Integer |  |  |
| river | Varchar(25) |  |  |
| description | Varchar(200) |  |  |

Таблица гидропосты (stations)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| station\_id | Integer |  |  |
| river\_id | Integer |  |  |
| station | Varchar(30) |  |  |
| method\_id | Smallint |  |  |
| description | Varchar(200) |  |  |

Таблица методики (methods)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| method\_id | Smallint (smallserial) |  |  |
| method | Varchar(100) |  |  |
| description | Varchar(200) |  |  |

Таблица предикторы (predictors\_groups)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| predictors \_id | Smallint (smallserial) |  |  |
| method\_id | Smallint |  |  |
| group\_n | Smallint |  | Номер группы предикторов в методике |
| predictors | Varchar(100) |  | Список предикторов |

Таблица алгоритмы (algorithms)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| algorithm\_id | Smallint (smallserial) |  |  |
| algorithm | Varchar(50) |  |  |
| linear | boolean |  |  |
| description | Varchar(200) |  |  |

Таблица моделей множественной регрессии (models)

|  |  |  |
| --- | --- | --- |
| Атрибут | Тип | Комментарий |
| model\_id | Integer (serial) | Уникальный ID обученной модели |
| algorithm\_id | Smallint |  |
| station\_id | Integer | ID гидропоста |
| predictors\_id | Smallint | ID группы предикторов, используемой при обучении модели множественной регрессии согласно методике расчета |
| Forecast\_year | Smallint | Год, на который рассчитан прогноз |
| Model\_file | Varchar(100) | Имя файла обученной модели на бэкенде |
| Group\_n | Smallint | Номер группы предикторов в методике |
| Predictors | Varchar(100) | Список предикторов в модели после обучения (строка) |
| Equations | Varchar(300) | Линейное уравнение множественной регрессии (строка) модели или надпись «Непараметрическая модель» |
| Dataset\_name | varchar | Строка «Река-Гидропост», например, «Неман-Белица» |
| algorithm | Varchar(50) | Наименование метода для обучения соответствующей модели множественной регрессии, например «LinearRegression» |
| Sigma | Numeric(5, 2) |  |
| Sigma\_t | Numeric(5, 2) |  |
| Delta\_dop | Numeric(5, 2) |  |
| Delta\_dop\_t | Numeric(5, 2) |  |
| Pm | Numeric(5, 2) |  |
| Pm\_t | Numeric(5, 2) |  |
| S | Numeric(5, 2) |  |
| S\_t | Numeric(5, 2) |  |
| Criterion | Numeric(10, 4) |  |
| Criterion\_t | Numeric(10, 4) |  |
| Correlation | Numeric(5, 4) |  |
| Correlation\_t | Numeric(5, 4) |  |
| R2 | Numeric(10, 4) |  |
| R2\_t | Numeric(10, 4) |  |
| Sigma\_f | Numeric(5, 2) |  |
| Delta\_dop\_f | Numeric(5, 2) |  |
| Pm\_f | Numeric(5, 2) |  |
| S\_f | Numeric(5, 2) |  |
| Criterion\_f | Numeric(10, 4) |  |
| Correlation\_f | Numeric(5, 4) |  |
| R2\_f | Numeric(10, 4) |  |

Таблица прогнозы (forecasts)

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| forecast\_id | Integer (serial) |  |  |
| forecast\_year | Smallint |  |  |
| model\_id | Integer |  |  |
| station\_id | Integer |  |  |
| predictors\_id | Smallint |  |  |
| h\_max | Smallint |  |  |
|  |  |  |  |
|  |  |  |  |

Таблица Наблюдения (observations) - наборы для обучения ММР

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| observation\_id | Integer (serial) |  |  |
| station\_id | Integer |  |  |
| obs\_year | Smallint |  |  |
| d\_0 | date |  | Дата перехода Т воздуха через 0 осенью |
| S\_2802 | Smallint |  | Снегозапасы на 28 февраля, мм |
| Smax | Smallint |  | Максимальные снегозапасы в мм |
| d\_smax | date |  | Дата формирования максимальных снегозапасов |
| H\_2802 | Smallint |  | Уровень воды над 0 поста на 28 февралясм, |
| H\_max | Smallint |  | Максимальный уровень воды над 0 поста, см |
| d\_hmax | date |  | Дата формирования максимального уровня воды над 0 поста |
| X | Smallint |  | Осадки от даты перехода температуры воздуха через 0 до даты формирования максимальных снегозапасов, мм |
| X1 | Smallint |  | Осадки от даты формирования максимальных снегозапасов до даты формирования максимального уровня воды над 0 поста, мм |
| X2 | Smallint |  | Осадки  от 01.03 до даты формирования максимального уровня воды над 0 поста, мм |
| X3 | Smallint |  | Осадки за 3 месяца до даты перехода температуры воздуха через 0, мм |
| Xs | Smallint |  | Осадки от даты перехода температуры воздуха через 0 до 01.03, мм |
| L\_2802 | Smallint |  | Глубина промерзания почвы на 28.02,  cм |
| L\_max | Smallint |  | Максимальная глубина промерзания почвы, см |
| d\_lmax | date |  | Дата формирования максимальной глубины промерзания почвы |
| Q12 | real |  | Среднемесячный расход воды в декабре, м3/с |
| Q01 | real |  | Среднемесячный расход воды в январе, м3/с |
| Q02 | real |  | Среднемесячный расход воды в феврале, м3/с |
| W12 | real |  | Объём стока в декабре, млн. м3 |
| W01 | real |  | Объём стока в январе, млн. м3 |
| W02 | real |  | Объём стока в феврале, млн. м3 |
| Y12 | real |  | Поверхностный  слой стока  в декабре,  мм |
| Y01 | real |  | Поверхностный  слой стока  в январе,  мм |
| Y02 | real |  | Поверхностный  слой стока  в феврале,  мм |
| Y\_sum | real |  | Поверхностный  слой стока  за зиму,  мм |

Таблица Данные для прогноза (predict\_data) - наборы для выполнения прогноза по новым данным.

|  |  |  |  |
| --- | --- | --- | --- |
| Атрибут | Тип | Длина |  |
| predict\_data \_id | Integer (serial) |  |  |
| station\_id | Integer |  |  |
| obs\_year | Smallint |  |  |
| d\_0 | date |  | Дата перехода Т воздуха через 0 осенью |
| S\_2802 | Smallint |  | Снегозапасы на 28 февраля, мм |
| Smax | Smallint |  | Максимальные снегозапасы в мм |
| d\_smax | date |  | Дата формирования максимальных снегозапасов |
| H\_2802 | Smallint |  | Уровень воды над 0 поста на 28 февралясм, |
| X | Smallint |  | Осадки от даты перехода температуры воздуха через 0 до даты формирования максимальных снегозапасов, мм |
| X1 | Smallint |  | Осадки от даты формирования максимальных снегозапасов до даты формирования максимального уровня воды над 0 поста, мм |
| X2 | Smallint |  | Осадки  от 01.03 до даты формирования максимального уровня воды над 0 поста, мм |
| X3 | Smallint |  | Осадки за 3 месяца до даты перехода температуры воздуха через 0, мм |
| Xs | Smallint |  | Осадки от даты перехода температуры воздуха через 0 до 01.03, мм |
| L\_2802 | Smallint |  | Глубина промерзания почвы на 28.02,  cм |
| L\_max | Smallint |  | Максимальная глубина промерзания почвы, см |
| d\_lmax | date |  | Дата формирования максимальной глубины промерзания почвы |
| Q12 | real |  | Среднемесячный расход воды в декабре, м3/с |
| Q01 | real |  | Среднемесячный расход воды в январе, м3/с |
| Q02 | real |  | Среднемесячный расход воды в феврале, м3/с |
| W12 | real |  | Объём стока в декабре, млн. м3 |
| W01 | real |  | Объём стока в январе, млн. м3 |
| W02 | real |  | Объём стока в феврале, млн. м3 |
| Y12 | real |  | Поверхностный  слой стока  в декабре,  мм |
| Y01 | real |  | Поверхностный  слой стока  в январе,  мм |
| Y02 | real |  | Поверхностный  слой стока  в феврале,  мм |
| Y\_sum | real |  | Поверхностный  слой стока  за зиму,  мм |

**Пользовательский интерфейс**

Создать выпадающее меню «Справочники» со следующими пунктами:

1. Реки
2. Гидропосты
3. Методики
4. Предикторы
5. Алгоритмы

Должны быть разработаны страницы интерфейса для каждого справочника в соответствии с их структурой (этап первый – вывод информации для просмотра, этап второй – предусмотреть возможность редактирования справочников (изменения, добавления, удаления).

Должны быть разработаны следующие страницы интерфейса с формами ввода-вывода информации:

1. «Наблюдения» - исходные данные многолетних наблюдений для обучения моделей множественной регрессии в разрезе каждого гидропоста в виде таблицы. Таблица должна предусматривать вывод данных на экран, ввод, корректировку, удаление данных по каждой строке наблюдений и иметь печатную форму (таблица ХХХ). Таблица Сортируется по колонке «Год», первой колонкой таблицы должен быть номер по порядку. Структура таблица в пользовательском интерфейсе должна соответствовать таблице базы данных observations и выводиться на экран по каждому гидропосту (поле station\_id таблицы stations).
2. «Редактирование наблюдения» - форма редактирования строки наблюдения по каждому году. Состав и порядок атрибутов для каждого гидропоста определяется таблицей БД predictors\_groups по ключевому полю method\_id и полю group\_n == 0.
3. «Данные для прогноза» – исходные данные для вычисления прогноза на указанный год в разрезе каждого гидропоста. Форма должна предусматривать ввод, корректировку, удаление данных, вывод данных на экран и иметь печатную форму (таблица ХХХ). Структура таблица в пользовательском интерфейсе должна соответствовать таблице базы данных predict\_data и выводиться на экран по каждому гидропосту (поле station\_id таблицы stations).
4. «Редактирование данных для прогноза» - форма редактирования строки наблюдения по каждому году. Состав и порядок атрибутов для каждого гидропоста определяется таблицей БД predictors\_groups по ключевому полю method\_id и полю group\_n == 0. При этом поля H\_max и d\_hmax не заполняются и в форме не отображаются.
5. «Оценка качества моделей МР» - оценки качества моделей множественной регрессии - статистические данные качества обученных моделей множественной регрессии в разрезе каждого гидропоста по совокупности обученных для этого гидропоста моделей в виде таблицы. Таблица должна предусматривать вывод данных на экран и иметь печатную форму (таблица ХХХ). Данные для таблицы должны извлекаться из таблицы models. Обязательными атрибутами для вывода в таблицу являются: Criterion, Correlation, Pm, Method, Predictors, Equations. Пример таблицы и печатной формы приведен в файлах <НаборДанных>-Оценка.xlsx (pk-ogo\source\_data\2024\Проверочные прогнозы)
6. «Проверочные прогнозы» – расчетные данные для проверки корректности алгоритмов и результатов обучения моделей множественной регрессии по каждой обученной модели и в разрезе каждого гидропоста. Форма должна предусматривать вывод данных на экран и иметь печатную форму. Форма рассчитывается в соответствии с методикой, описанной в пункте ППП.
7. «Прогнозы» – результаты вычисления прогнозов максимального уровня воды в разрезе каждого гидропоста, на заданный год, по каждой обученной модели.

В пользовательском интерфейсе должен быть реализован эргономичный (удобный для пользователя) механизм доступа к указанным выше данным, например:

1) на экран выводится список гидропостов в виде ссылок;

2) напротив каждого гидропоста размещен набор следующих ссылок:

- «Наблюдения»

- «Данные для прогноза»

- «Оценки качества моделей МР»

3) при щелчке на ссылку «Наблюдения» открывается страница «Наблюдения» с данными многолетних наблюдений по данному гидропосту в виде таблицы;

3.1) при выборе режима редактирования (ввода) строки данных открывается страница «Редактирование наблюдения» с формой редактирования (ввода) данных наблюдения за выбранный год;

3.2) при выборе режима удаления строки, строка данных удаляется;

4) при щелчке на ссылку «Данные для прогноза» открывается страница «Данные для прогноза» с данными для прогноза по данному гидропосту на указанные годы в виде таблицы;

4.1) при выборе режима редактирования (ввода) строки данных открывается страница «Редактирование данных для прогноза» с формой редактирования (ввода) строки данных для прогноза на заданный год;

4.2) при выборе режима удаления строки, строка данных удаляется;

5) при щелчке на ссылку «Оценка качества моделей МР» открывается страница «Оценка качества моделей МР» с возможностью выбора года и таблицей, содержащей перечень обученных моделей множественной регрессии и соответствующим им статистическим данным качества их обучения (оценок) по соответствующему гидропосту за выбранный год. Строки моделей являются ссылками.

5.1) при щелчке по строке модели (ссылке) открывается страница «Проверочный прогноз», содержащая таблицы с данными проверочного прогноза выбранной модели множественной регрессии. Данные проверочного прогноза являются расчетными. Форма проверочного прогноза приведена в таблице ХХХ. Алгоритм расчета в пункте ППП.

«Наблюдения»

«Редактирование наблюдения»

«Данные для прогноза»

«Редактирование данных для прогноза»

«Оценка качества моделей МР»

«Проверочный прогноз»

«Прогноз»

**Требования к ПО на стороне сервера (бэкенд).**

Для работы прогностических задач на стороне сервера должно быть установлено следующее ПО:

python 3.10.11

scikit-learn 1.2.2

numpy 1.24.3

scipy 1.10.1

matplotlib 3.7.1