Database Outline & Schema

nechanie

January 2023

Overview

In the design of this database, the following use case requirements have been defined. The database will be used to store the run-time analytics of the deep learning model. This will include entries for each execution of the model. Additionally, the database will be used to store the output and input of each model execution. The output will be a very large dataset representing the vector returned by the model, and the input will be a variably large sized representation of the initial list of input items. It is expected that the tables within the database will store hundreds of thousands of rows. Due to this expectation, the data within these rows will be optimized to contain mainly reference data, minimizing the space required by each row. Lastly, this database will store the model information from each model execution. This will include, at minimum, a single row for each model execution.

Database Outline

- Exec_Info:∼ Stores the data related to program execution.

    - exec_id:∼ Unique auto-incrementing execution identifier.
    - exec_start:∼ Execution start time.
    - exec_stop:∼ Execution stop time.

- Exec_Input:∼ Stores the input data related to program execution.

    - input_id:∼ Unique auto-incrementing input identifier.
    - exec_id:∼ Foreign Key connecting input with its corresponding execution.
    - person_id:∼ Identifier for person in input item, id from external database used.

    Foreign Key from Exec_Info 1:M relationship

- Exec_Output:∼ Stores the output data related to program execution.

- output_id:∼ Unique auto-incrementing output identifier.
- exec_id:∼ Foreign Key connecting output with its corresponding execution.
- person_id:∼ Identifier for person in output item, id from external database used.
- k_value:∼ Similarity value produced by model with respect to input.

Foreign Key from Exec_Info 1:M relationship

- Exec_Model:∼ Stores the model information related to program execution.
    - model_id:∼ Unique auto-incrementing model identifier.
    - exec_id:∼ Foreign Key connecting output with its corresponding execution.
    - TODO:∼ hyperparameter values for current model.
    - mod_start:∼ Model run-time start time.
    - mod_stop:∼ Model run-time stop time.

Foreign Key from Exec_Info 1:1 relationship

ER Diagram
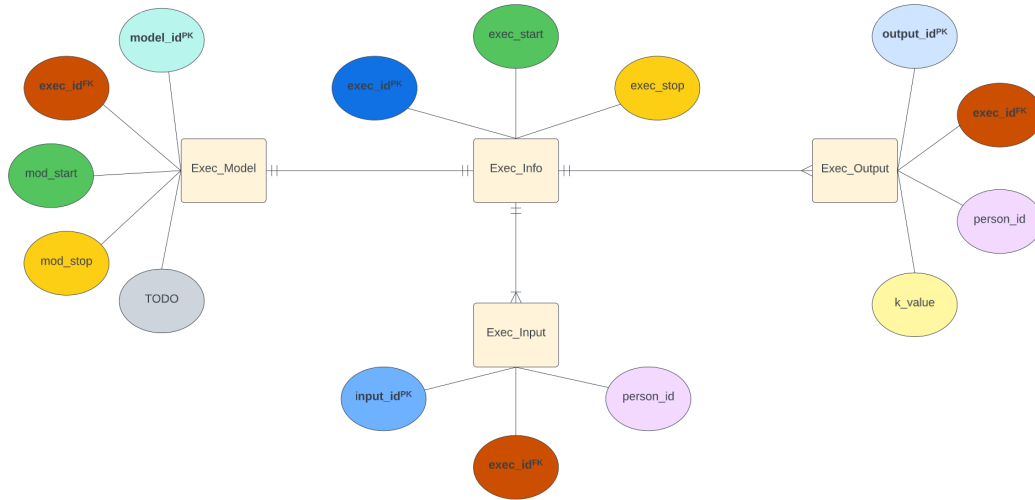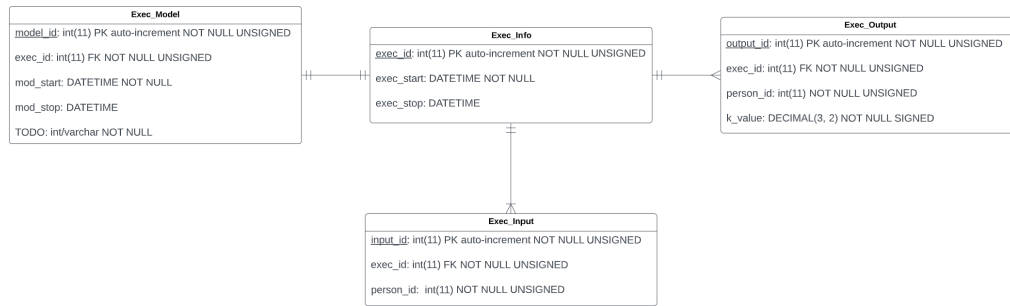

Figure 1: Database ERD

Schema

Figure 2: Database Schema

- Exec_Info:

    - exec_id:

        * int(11)
        * auto-increment
        * NOT NULL
        * UNSIGNED
        * PRIMARY KEY

    - exec_start:

        * DATETIME
        * NOT NULL

    - exec_stop:

        * DATETIME

- Exec_Input:

    - input_id:

        * int(11)
        * auto-increment
        * NOT NULL
        * UNSIGNED
        * PRIMARY KEY

    - exec_id:

        * int(11)
        * FOREIGN KEY

3

* ∗ NOT NULL
  * ∗ UNSIGNED
  - – person_id:
    * ∗ int(11)
    * ∗ NOT NULL
    * ∗ UNSIGNED

- • Exec_Output:
  - – output_id:
    * ∗ int(11)
    * ∗ auto-increment
    * ∗ NOT NULL
    * ∗ UNSIGNED
    * ∗ PRIMARY KEY
  - – exec_id:
    * ∗ int(11)
    * ∗ FOREIGN KEY
    * ∗ NOT NULL
    * ∗ UNSIGNED
  - – person_id:
    * ∗ int(11)
    * ∗ NOT NULL
    * ∗ UNSIGNED
  - – k_value:
    * ∗ DECIMAL(3, 2)
    * ∗ NOT NULL
    * ∗ SIGNED

- • Exec_Model:
  - – model_id:
    * ∗ int(11)
    * ∗ auto-increment
    * ∗ NOT NULL
    * ∗ UNSIGNED
    * ∗ PRIMARY KEY
  - – exec_id:
    * ∗ int(11)
    * ∗ FOREIGN KEY

- * NOT NULL
- * UNSIGNED
- mod_start:
  - * DATETIME
  - * NOT NULL
- mod_stop:
  - * DATETIME
- TODO:
  - * int/VARCHAR
  - * NOT NULL