**Overview**

The **Department of Computer Science and Engineering (CSE)** at KUET proudly organizes its **first-ever Datathon**, an exhilarating opportunity for data enthusiasts to showcase their skills and creativity in the rapidly growing field of data science and analytics. This landmark event aims to promote data-driven innovation, provide hands-on experience with real-world datasets, and inspire participants to excel in solving complex data challenges.

**Key Sponsors and Partners**

* **Title Sponsor:** **Blockstak**
* **Co-Sponsor:** **Daamdekhi.com**
* **Technical Partner:** **NeuralFrameAI**  
    
    
  Organized by KUET’s CSE Department, this Datathon marks a significant milestone in encouraging competitive learning and nurturing young talents in data science. Whether you are an experienced or a beginner looking to explore the world of data, this competition provides an incredible platform to learn, collaborate, and shine.  
  Join us in making history at KUET’s **first-ever Datathon** and unlock your potential in the fascinating realm of data science!

**Start**

17 hours ago

**Close**

a day to go

**Description**

link

keyboard\_arrow\_up

**What to Expect:**

* **Online Round:** This is the online round of the Datathon event at Bitfest 2025.
* **Dataset:** You'll be provided with a compact, real-world dataset designed to test your analytical and predictive modeling skills.
* **Timeline:** The competition will run from **26th December 6:00 PM** to **28th December 6:00 PM**.
* **Scoring:** Submissions will be evaluated using the **Mean Squared Error (MSE)** score.  
  We encourage all participants to put their best foot forward and showcase their data science prowess!  
  **Best of luck!**

**Evaluation**

link

keyboard\_arrow\_up

**Evaluation Metric**

The performance of submissions in this competition will be evaluated using **Mean Squared Error (MSE)**. MSE measures the average squared difference between predicted and actual values, with lower values indicating better model performance. The formula for MSE is:

MSE=1n∑i=1n(yi−y^i)2MSE=1n∑i=1n(yi−y^i)2

Where

* ***n***: the number of data points
* ***yᵢ***: the actual value
* ***ŷᵢ***: the predicted value

**Leaderboard**

The leaderboard will be determined based on the **MSE** score of each submission. **Submissions with lower MSE values will be ranked higher.** The final rankings reflect the ability of the models to minimize the prediction error.

**Submission Limits**

link

keyboard\_arrow\_up

a. Your team may submit a maximum of **ten (10) Submissions** per day.  
b. Your team may select up to **four (4) Final Submissions** for judging.

**Scoring Guidelines**

link

keyboard\_arrow\_up

Participants are required to use Kaggle notebooks for this competition. Ensure that your code is well-commented and incorporates appropriate visualizations to support your findings. Once completed, the notebook must be **submitted** to the host for evaluation, as the overall quality of both the code and documentation will be integral to the scoring process.

**Evaluation Criteria**

The competition will be assessed based on the following criteria:

**Milestones**

* **Data Manipulation:** Effective preprocessing and handling of the dataset.
* **Data Analysis:** Ability to extract meaningful insights from the data.
* **Using Methods:** Correct and innovative application of algorithms or techniques.

**Online Round Scoring**

The score from the online round for each team will be calculated using the following formula:

FinalScore=80×MSEWorst−MSESubmissionMSEWorst−MSEBestFinalScore=80×MSEWorst−MSESubmissionMSEWorst−MSEBest

**Explanation of Terms:**

* **MSE\_Worst**: The highest Mean Squared Error (MSE) value observed across all submissions of all team
* **MSE\_Best**: The lowest Mean Squared Error (MSE) value observed across all submissions of all team
* **MSE\_Submission**: The MSE of the model submitted by the team.

**Presentation**

* Clear and concise explanation of the solution.

**Documentation**

* Properly structured and well-organized documentation of the work, embedded in the same notebook as the code. The same notebook should include both the code and the documentation (text sections within the notebook).

**Scoring Breakdown**

* **Online Round Score (Leaderboard):** 80% of the total marks.
* **Onsite Presentation:** 15% of the total marks, further divided into:
  + **Presentation Quality:** 5%.
  + **Question-and-Answer Session with Judges:** 10%.
* **Documentation:** 5% of the total marks.

This comprehensive structure ensures both the technical and presentational aspects of your work are evaluated effectively.

ock

**Dataset Description**

This competition involves predicting the **matched\_score**, a continuous variable indicating the match between a candidate's profile and job requirements.

**Files**

1. **train.csv**: The training dataset containing features and the target variable matched\_score.
2. **test.csv**: The test dataset for predictions, excluding the matched\_score column, with a unique ID for each row.
3. **sample\_submission.csv**: A template for submissions, containing ID and placeholder matched\_score values to be replaced with predictions.

**Dataset Format**

Participants will encounter features in various formats, including integers, floating-point numbers, text fields and lists.

**Prediction Task**

Participants must predict the **matched\_score** for rows in test.csv and submit predictions using the format provided in sample\_submission.csv.

**Files**

3 files

**Size**

17.01 MB

**Type**

csv

**License**

[MIT](https://www.mit.edu/~amini/LICENSE.md)

**sample\_submission.csv**(16.09 kB)

get\_app

fullscreen

chevron\_right

DetailCompactColumn

2 of 2 columns

keyboard\_arrow\_down

vpn\_keyIDsortgrid\_3x3matched\_scoresort

| **Label** | **Count** |
| --- | --- |
| 1.00 - 191.80 | 191 |
| 191.80 - 382.60 | 191 |
| 382.60 - 573.40 | 191 |
| 573.40 - 764.20 | 191 |
| 764.20 - 955.00 | 190 |
| 955.00 - 1145.80 | 191 |
| 1145.80 - 1336.60 | 191 |
| 1336.60 - 1527.40 | 191 |
| 1527.40 - 1718.20 | 191 |
| 1718.20 - 1909.00 | 191 |

1

1909

| **Label** | **Count** |
| --- | --- |
| 0.50 - 0.50 | 1,909 |

0.5

0.5

10.520.530.540.550.560.570.580.590.5100.5110.5120.5130.5140.5150.5160.5170.5180.5190.5200.5210.5220.5230.5240.5250.5260.5270.5280.5290.5300.5310.5320.5330.5340.5350.5360.5370.5380.5390.5400.5410.5420.5430.5440.5450.5460.5470.5480.5490.5500.5

**Data Explorer**

17.01 MB

* calendar\_view\_week

sample\_submission.csv

* calendar\_view\_week

test.csv

* calendar\_view\_week

train.csv

**Summary**

arrow\_right

folder

3 files

arrow\_right

calendar\_view\_week

72 columns

get\_app**Download All**

Download data

navigate\_nextminimize

content\_copyhelp

Download data

text\_snippet

**Metadata**

