

# **Term Projects:**

## **Applying EC techniques to find optimized DL models for suggested projects or your own project**

Fall 2025

*V5, Updated on Nov. 6*

# Suggested Project 1: Find Defective Pallets

- Jayshop: Empowering disenfranchised and disabled workers through gainful employment
- Video: <https://photos.app.goo.gl/7MLvciFa5YoqARHXA>
- **Non Disclosure Agreement (NDA) is needed**
- Need GPU

## Student Non-Disclosure Agreement (NDA) for Class Dataset

### Parties:

- **Disclosing Party:** Professor Dr. Chan-Jin "CJ" Chung, on behalf of LTU ("University").
- **Receiving Party:** \_\_\_\_\_, Student ID: \_\_\_\_\_ ("Student").

**Date:** October \_\_, 2025

### 1. Purpose

For the sole purpose of completing coursework and assignments ("Permitted Purpose") for the class MCS5993 EC and DL ("Class") during the Fall 2025 semester, the University will provide the Student with access to certain Confidential Information.

### 2. Definition of Confidential Information

"Confidential Information" refers to the proprietary image dataset known as jayshop's automotive pallet image datasets (the "Dataset"), including but not limited to:

- All images, video files, and visual data.
- All associated annotations, labels, and metadata.
- Any subsets, derivatives, or analysis results created by the Student that include or display the Dataset.

### 3. Student's Obligations

The Student hereby agrees:

- **Use Limitation:** To use the Confidential Information solely for the Permitted Purpose and for no other personal or commercial reason.
- **Non-Disclosure:** To hold the Confidential Information in strict confidence. The Student will not disclose, publish, distribute, post online (e.g., GitHub, social media, personal websites), or otherwise share the Confidential Information with any other person or entity.
- **Security:** To use all reasonable measures to protect the Dataset from unauthorized access, copying, or disclosure.
- **No Copying:** Not to transfer or copy the Confidential Information to any person or device not directly involved in completing the Permitted Purpose (e.g., no sharing with other students unless part of a group assignment where all members have signed this NDA).

### 4. Term and Data Destruction

The Student's access to the Confidential Information shall terminate upon the submission of final grades for the Class. Upon termination, the Student must promptly and permanently delete and

destroy all copies of the Confidential Information (digital or otherwise) from all personal devices, cloud storage, and storage media.

### 5. Breach of Agreement

The Student's obligation of non-disclosure shall survive the conclusion of the Class. The Student acknowledges that any breach of this Agreement constitutes a violation of academic integrity and may subject the Student to disciplinary action under the University's Code of Student Conduct, up to and including a failing grade for the Class. The University also reserves the right to seek any available legal remedies for damages caused by the breach.

### 6. Acknowledgment

I, the Student, have read this entire Agreement, understand its terms, and agree to be bound by them as a condition of receiving access to the Confidential Information.

---

### STUDENT (Receiving Party):

Signature: \_\_\_\_\_

Printed Name: \_\_\_\_\_

Date: \_\_\_\_\_

### PROFESSOR (Disclosing Party):

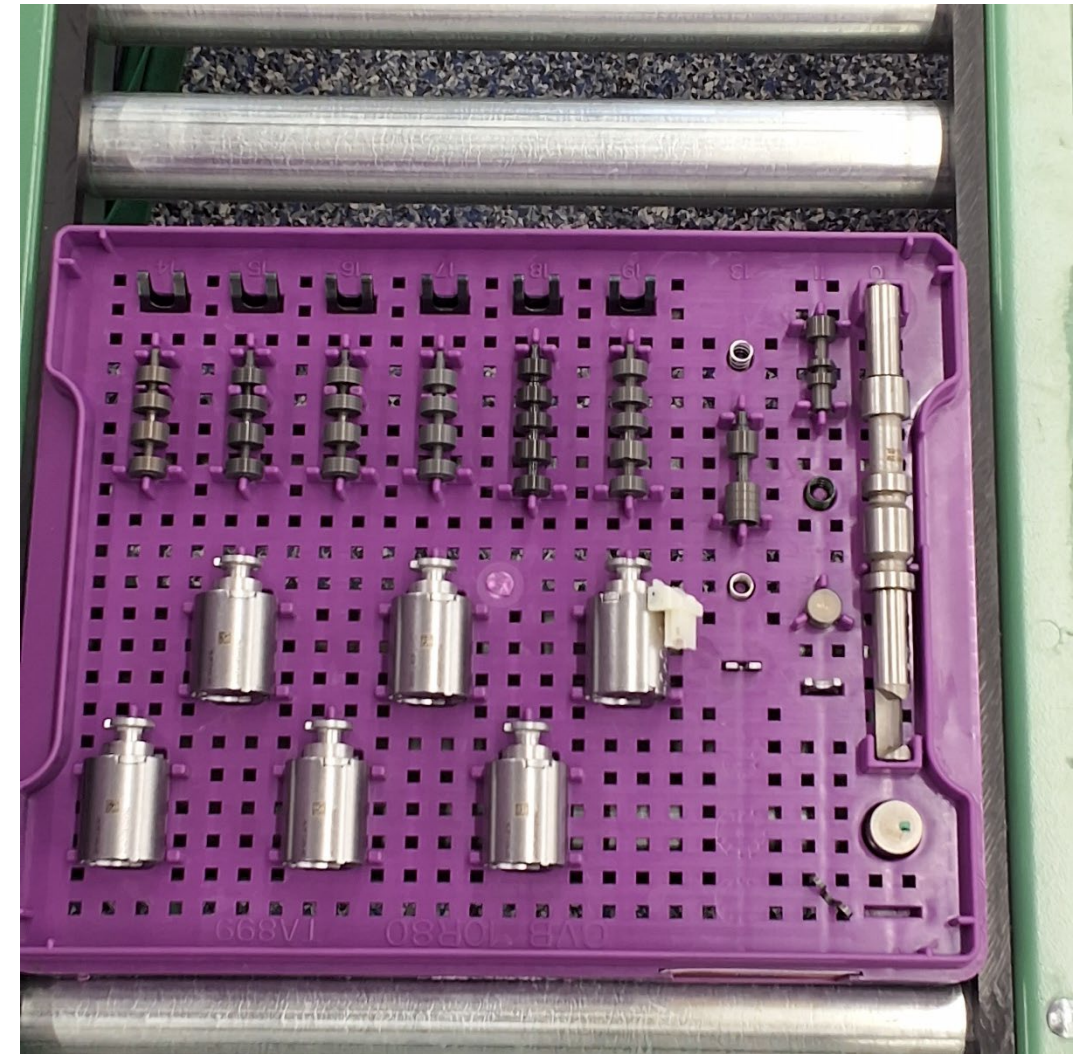
Signature: \_\_\_\_\_

Printed Name: Chan-Jin "CJ" Chung

Date: \_\_\_\_\_

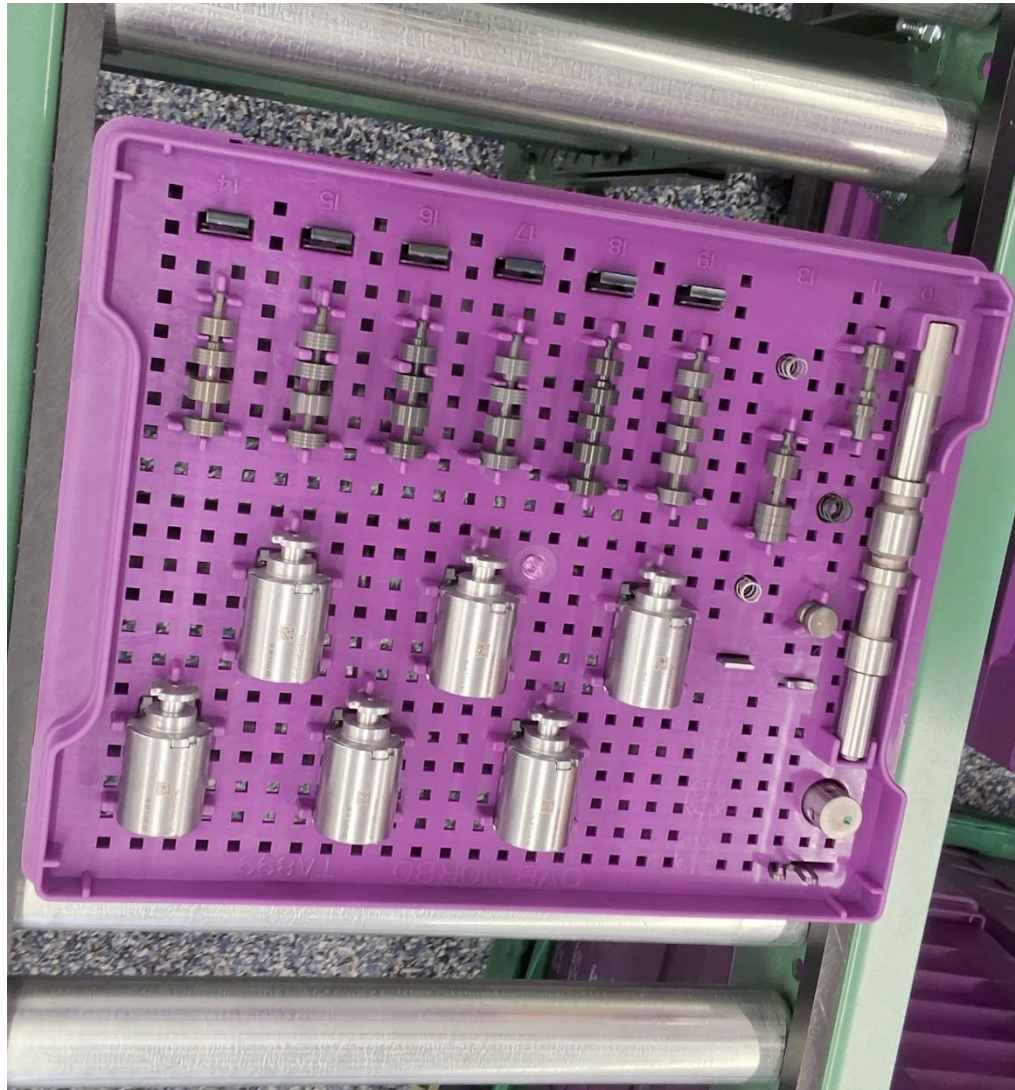


Good



Bad



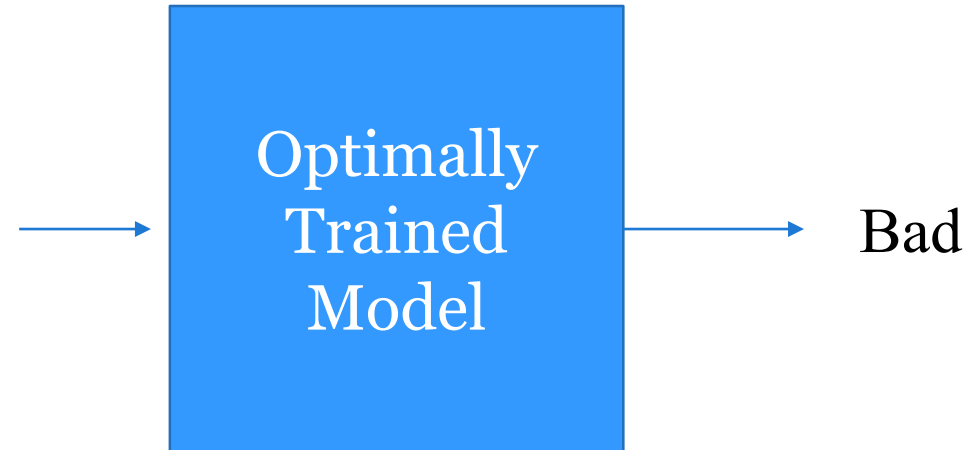


Good



Bad

# Test



- ES(1+1) with 1/5 rule
- Another Evolutionary Algorithm
- Random Search

# Accessible only for Abdul

```
=====
Dataset root: /content/drive/MyDrive/DL_data/jayshop/pallets
=====

[test]
  class: bad           31 images
  class: good          20 images
  -----
  Total images in test : 51

[train]
  class: bad           95 images
  class: good          67 images
  -----
  Total images in train : 162

[validation]
  class: bad           31 images
  class: good          20 images
  -----
  Total images in validation: 51
=====
```



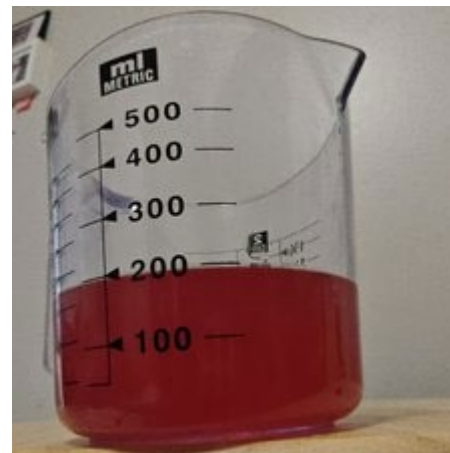
# Suggested Project 2: Regression Problem - Estimate the volume of a measuring cup in ML



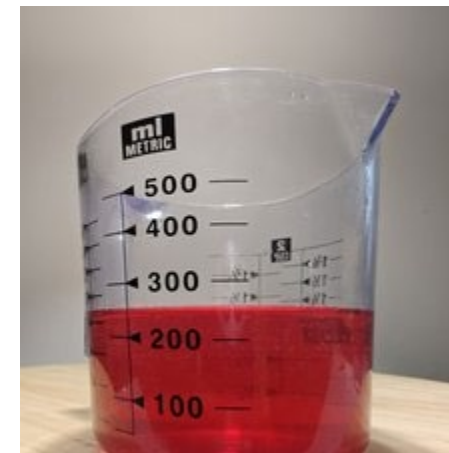
100 ml



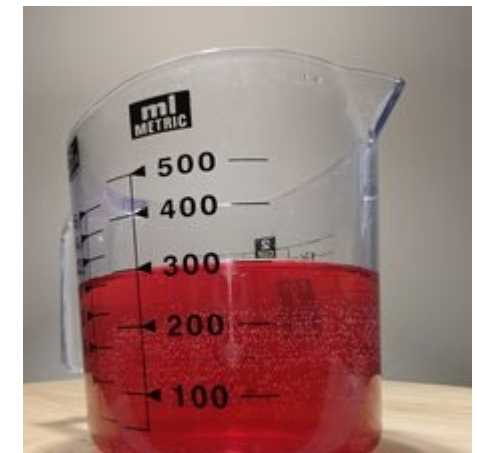
150 ml



200 ml



250 ml



300 ml



350 ml



400 ml



450 ml

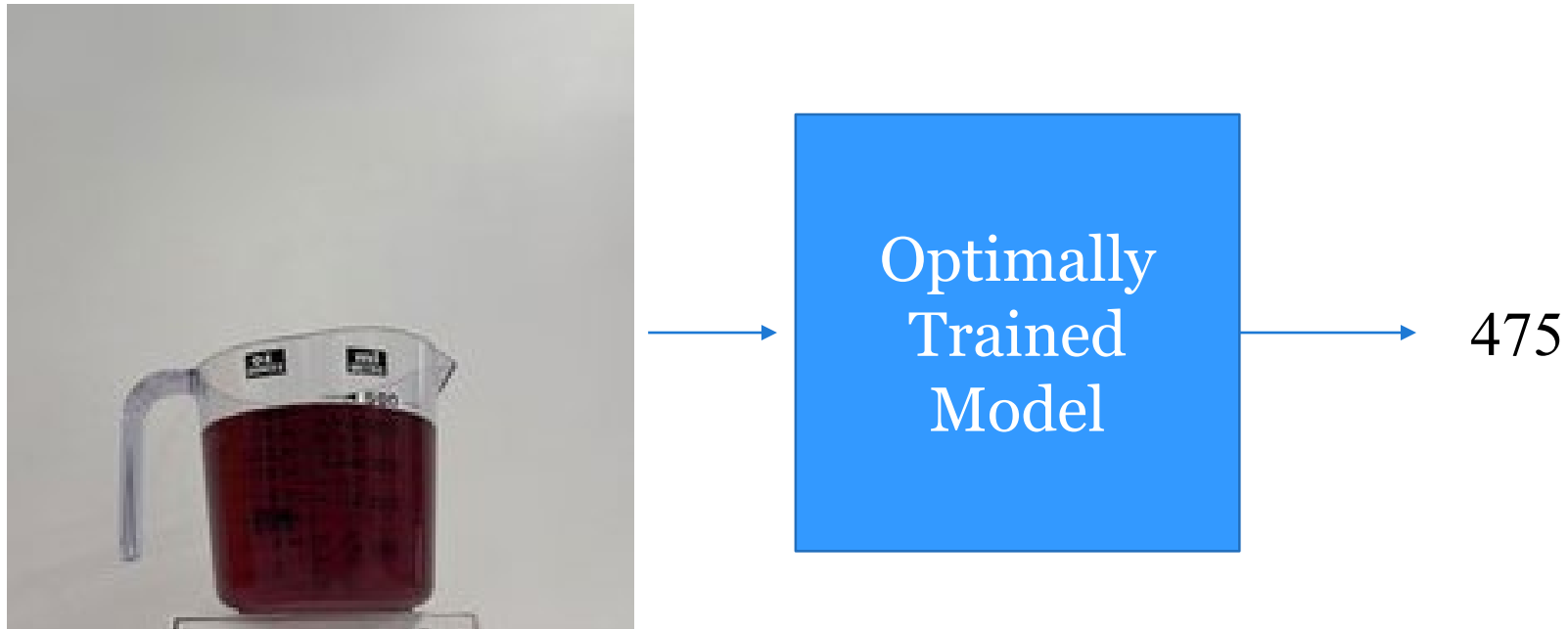


500 ml

Need GPU



# Test



- ES(1+1) with 1/5 rule
- Another Evolutionary Algorithm
- Random Search

# Datasets

- No separate validation folder
- Split your training dataset into training and validation sets without creating a separate directory using the **validation\_split** and subset arguments in **tf.keras.utils.image\_dataset\_from\_directory**.

## Dataset Link:

[https://drive.google.com/drive/folders/1xSxWOkBX2ZanNMQN2neNeW\\_0f4PoWzkL?usp=drive\\_link](https://drive.google.com/drive/folders/1xSxWOkBX2ZanNMQN2neNeW_0f4PoWzkL?usp=drive_link)

```
=====
Dataset root: /content/drive/MyDrive/DL_data/mcup
=====

[test]
class: 105      21 images
class: 155      22 images
class: 205      44 images
class: 260      50 images
class: 310      35 images
class: 355      44 images
class: 410       9 images
class: 470      43 images
class: 505      43 images
-----
Total images in test      : 311

[train]
class: 100      43 images
class: 150      44 images
class: 200      44 images
class: 250      44 images
class: 300      44 images
class: 350      40 images
class: 400       4 images
class: 450      44 images
class: 500      44 images
-----
Total images in train     : 351
=====
```

Will provide more data later. Please start to test as is.

# Suggested Project 3: Leveraging EC for Your Own Image Classification or regression Problem

- Application must be unique. (Real-world project preferred)
- Required Model Optimization Algorithms
  - ES(1+1) with 1/5 rule
  - Another Evolutionary Algorithm such as using Cov. Matrix or Genetic Algorithms
  - Random Search



# Suggested Project 4: Time series forecasting with Multivariate Time Series Datasets – *Extension of HW5*

- ES(1+1) with 1/5 rule
- Another Evolutionary Algorithm such as GA and Cov. Mat.
- Random Search
- You may not need GPU
- **Important:** DO NOT use Lag Data in the training and test datasets. Use the way in Quiz #3.  
“TimeSeries\_Univari\_HPO.ipynb” file is on Canvas.

# Suggested Project 5: sequence-to-sequence (seq2seq) time series prediction using “Jena Weather Dataset”

- Seq2Seq time series prediction means using past time-series data (a sequence) to predict future time-series data (another sequence) — not just one next point, but several future steps.
- Multi-step horizons (Not just one step ahead, “seq2one”)
- Optimize the model using 2 EAs
- <https://www.kaggle.com/datasets/harishedison/jena-weather-dataset/data>

# Suggested Project 6 – Any research oriented project using EC and DL

- Must prepare a draft poster or paper at the end the semester



# Use Discussions to express your initial interest

- Choose one out of 6 suggested projects.
- Group project possible up to 2 members
- **Due by Sat. Oct 18**
- Final decision/approval must be done by **Wed. Oct 22**

# Project Evaluation

- Completed as planned?
- Functionality: Work as intended?
- User Interface (if applicable): Is the project user-friendly and accessible?
- Complexity: Is the project appropriately challenging for a senior-level?
- Code quality: Is the code clean, maintainable, and well-documented?  
Are best practices, standards followed?
- Size: not too small for a semester project?
- Others
  - Innovation? Does the project demonstrate creativity or a novel approach?
  - Scalability: Could the project handle growth in users or data volume?
  - Documentation (Poster, Papers)
  - Peer evaluation
  - ...

# Project Evaluation, Continued

- Expected final product: A poster for “LTU Research Day”, Friday, April 24, 2026
- Clearly specify the author of each result
- Clearly specify the author of each poster section/subsection
- Peer evaluation



# Current 10 Project Proposals (as of Oct 27)

## Any Changes?

Name	Project Title	Notes
Jacob Moore	Evolutionary Strategies for Machine Translation Prompt Engineering for Low-Resource Languages	
Abdul Kareem Ansari	Automated Detection of Defective Pallets Using Evolutionary Computation Optimized Deep Learning	CNN+DNN HPO
V Harsha Vardhan Yellela	Deep Food Image Classifier with Binary Expert Layer: A CNN-HPO Approach with ES(1+1)	Need datasource
Venkata Sayee Joginipally, Anirudh Komaragiri	Regression Problem - Estimate the volume of a measuring cup in ML	CNN+DNN HPO
Ravali Kamindla, Gouthami Vasam	Time series forecasting with Multivariate Time Series Datasets. (Extension of HW5 from Suggested Project 4)	
Renuka Jayvant Jagadale	Time series forecasting with Multivariate Time Series Datasets (Extension of HW5)	
✓ Fnu Mohammed Abdul Nafeh	Car Damage Cost Predictor Using Deep Learning Optimized with Evolutionary Algorithms	No dataset
Travis Bowman, Anthony Gabrail	Battery classification and segmentation	No mention about EC
✓ Deepak Goud Nalla	Hand Gesture Recognition using Deep Learning	No mention about EC
✓ Sathwika Kuppirala & Manikanta Jakkidi	Regression Problem: Estimate the volume of a measuring cup in ML(Suggested Project 2)	CNN+DNN HPO

OK, but too many classes

OK, but too big dataset

OK, U-Net HPO



# Course Plan, Continued

- HW6 – Class Competition. To be announced on Nov 5.
- 1<sup>st</sup> Project Presentation on **Mon. Nov 10**
  - Project goals and description
  - Dataset description
  - Any Test Results

} => Poster Abstract
- 2nd Project Presentation for (1 point) on **Wed. Nov 19**
  - Project goals and description
  - Dataset description
  - Any Test Results with basic EC algorithm (for HPO)
- **Fri Nov 21, Last Day to drop classes**

# Extra Credit Opportunity 1

- Do the project 5 - sequence-to-sequence (seq2seq) time series prediction using “Jena Weather Dataset”

# Extra Credit Opportunity 2

- Do some extra work for your chosen project

- 4<sup>th</sup> HPO algorithm
- Using “Keras Tuner” for HPO.

<https://www.geeksforgeeks.org/deep-learning/keras-tuner-for-hyperparameter-optimization/>

- **Random Search:** Samples hyperparameters randomly from the space; simple but potentially inefficient.
- **Bayesian Optimization:** Uses past evaluation results to guide the search toward promising regions. It uses probabilistic models (like Gaussian processes) to choose next parameters, aiming for efficiency.
- **Hyperband:** Uses early stopping and adaptive resource allocation to speed up the search by pruning bad trials early, effectively balancing exploration and exploitation.

- The extra work must be approved by the instructor



# Extra Credit Opportunity 3

- HW6 Class Competition
- Unknown optimization function will be used