

Configuration Manual

MSc Research Project
MSc in Cloud Computing

Ashwini Ashok Patil
Student ID: 20218745

School of Computing
National College of Ireland

Supervisor: Jitendra Kumar Sharma

**National College of Ireland
Project Submission Sheet
School of Computing**



Student Name:	Ashwini Ashok Patil
Student ID:	20218745
Programme:	MSc in Cloud Computing
Year:	2022
Module:	MSc Research Project
Supervisor:	Jitendra Kumar Sharma
Submission Due Date:	15/08/2022
Project Title:	Early Prediction of HDD Failures in the Cloud Using Interpretable AI Models
Word Count:	6000 aprx
Page Count:	2

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Ashwini Ashok Patil
Date:	15th August 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission , to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project , both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Ashwini Ashok Patil
20218745

1 Tools Used to implement the Project

- Frameworks for Machine Learning: sklearn, Interpretable AI toolkit
- Framework for Deep Learning: PyTorch
- Coding Language: Python
- Other Tools: Hadoop(for ETL), pandas, matplotlib, numpy, seaborn, Flask framework

2 Steps to work with the .zip file of Source Code

- Create python venv
- Use the requirements.txt to install necessary packages to setup the environment required to run the code: `pip install -r requirements.txt`
- LSTM models training code can be found in `bi_lstm_model` folder
- Steps to train model:
 - `cd bi_lstm_model`
 - `source init_env.sh`
 - configure model params in `configs/train_bilstm.config`
 - `python flows/train_bilstm.py run`
 - Best performing and results can be found in `artifacts/ckpts` folder
- RFC model training code can be found in `rfc_model` folder Steps to train model:
 - `python train_rfc.py`
 - Trained models and results can found in `model/` and `res/` folders respectively
- OCT model training code can be found in `oct_model` folder Steps to train model:
 - The OCT tree implementation depends on GUROBI solver. Please follow instructions on gurobi website to install and acquire license ¹

¹GUROBI:<https://www.gurobi.com/academia/academic-program-and-licenses/>

- python train_oct.py
- Trained models and results can found in model/ and res/ folders respectively
- The composite_model code can be found in composite_model.py Steps to evaluate composite model:
 - python composite_model.py
 - results can found in res/ folder
- The composite_model is integrated to webapp for demo. The source for the same can be found in hdd_health_monitor_app. Steps to run the app locally:
 - cd hdd_health_monitor_app
 - flask run
 - App can be accessed at url: localhost:5000
 - The app assumes that data fed for predictions is already scaled to mean 0 and std 1
 - Test files for app demo can be found in data/test_data_app folder

3 Section 3

- The Jupiter Notebook installation can be found here. ²
- This code makes use of python implementation of OCT tree by Bo Tang: ³
- Due to installation and license issues for interpretable AI package, it was decided to go with python implementation of OCT tree by Bo Tang.

References

²Jupyter Install:<https://jupyter.org/install>

³OCT code reference:https://github.com/LucasBoTang/Optimal_Classification_Trees