



MACQUARIE
University



Automatic Exam Marking Project

Key Task Description – Step 1

PMP Design

Wireframe Design

Using Chat GPT for Dataset Generation

GAN for Essay Dataset Generation (research and Code Analysis)

&

Essay Scoring (research and Code Analysis)

Task 1: Project Management Plan (PMP)

• Prepare PMP Document & Execution

Expected Task Output:

- Project Documentation Plan
 - Preparing Template for weekly reports
 - Document Versioning structure
 - Selecting a tools for document Management
- Project Monitoring
 - Prepare Detailed Gant Chart for upcoming Phases
 - Updating Project Gant Chart
 - Project Progress Report Template
- Project Risk Management
 - Prepare a template for reporting project risks weekly
 - Select a method to Analysis Risks
- Project Resource Management
 - Resource Requirement Analysis
 - Resource Allocation
- Project Stakeholder Management
 - Meeting Management
 - Feedback Coordination
 - Customer meeting

Task 2: Using Chat GPT for Data Source Generation

Expected Task Output:

- Using JSON to gather Data From Chat GPT by
- Recognizing Limitations
- Selecting the Sample Questions
- Preparing a Demo
- Plan for Data Generation

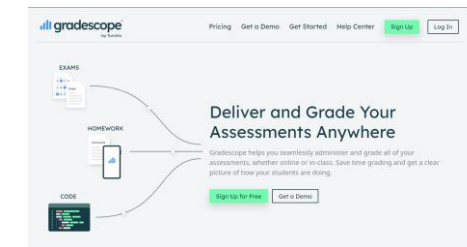
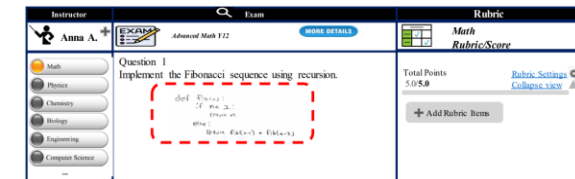
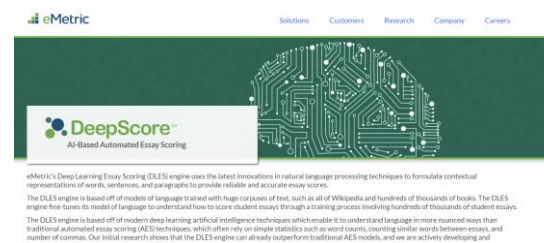
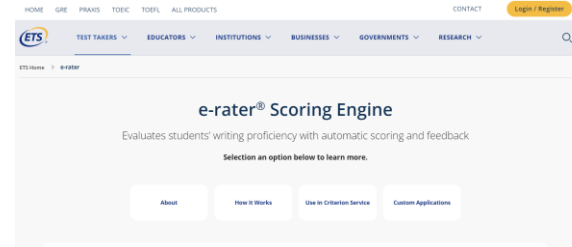
Task 3: Commercial Tools Analysis & Wireframes Design

Commercial Tools Analysis:

- Functionality
 - Different Roles and Users
 - Scope
 - UI Functions
- Data Sets
- Technology
 - Data Management
 - UI
 - ML Engines
- Method and approaches
 - Finding Related Paper
 - ML approach for Scoring
 - ML Approach for Data Set generation

Expected Task Output:

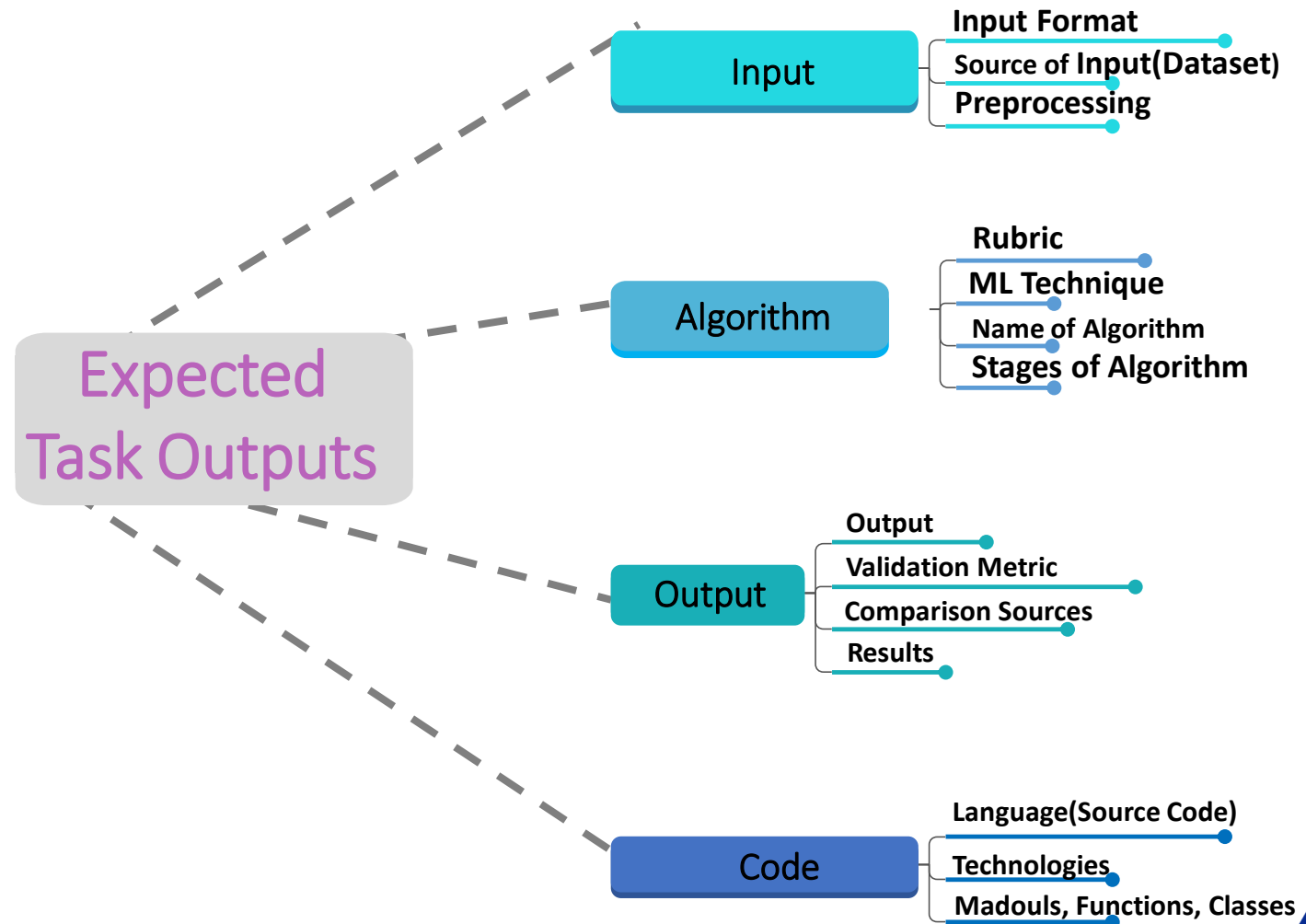
- Commercial Tools Functionality, UI and Scope Comparison
- Gap Analysis with moc up
- Design Wireframes



Task 4: Existing GAN For Generation Essay DataSet

Research & Code Analysis

Park, Y. H., Choi, Y. S., Park, C. Y., & Lee, K. J. (2022). EssayGAN: Essay Data Augmentation Based on Generative Adversarial Networks for Automated Essay Scoring. *Applied Sciences*, 12(12), 5803.

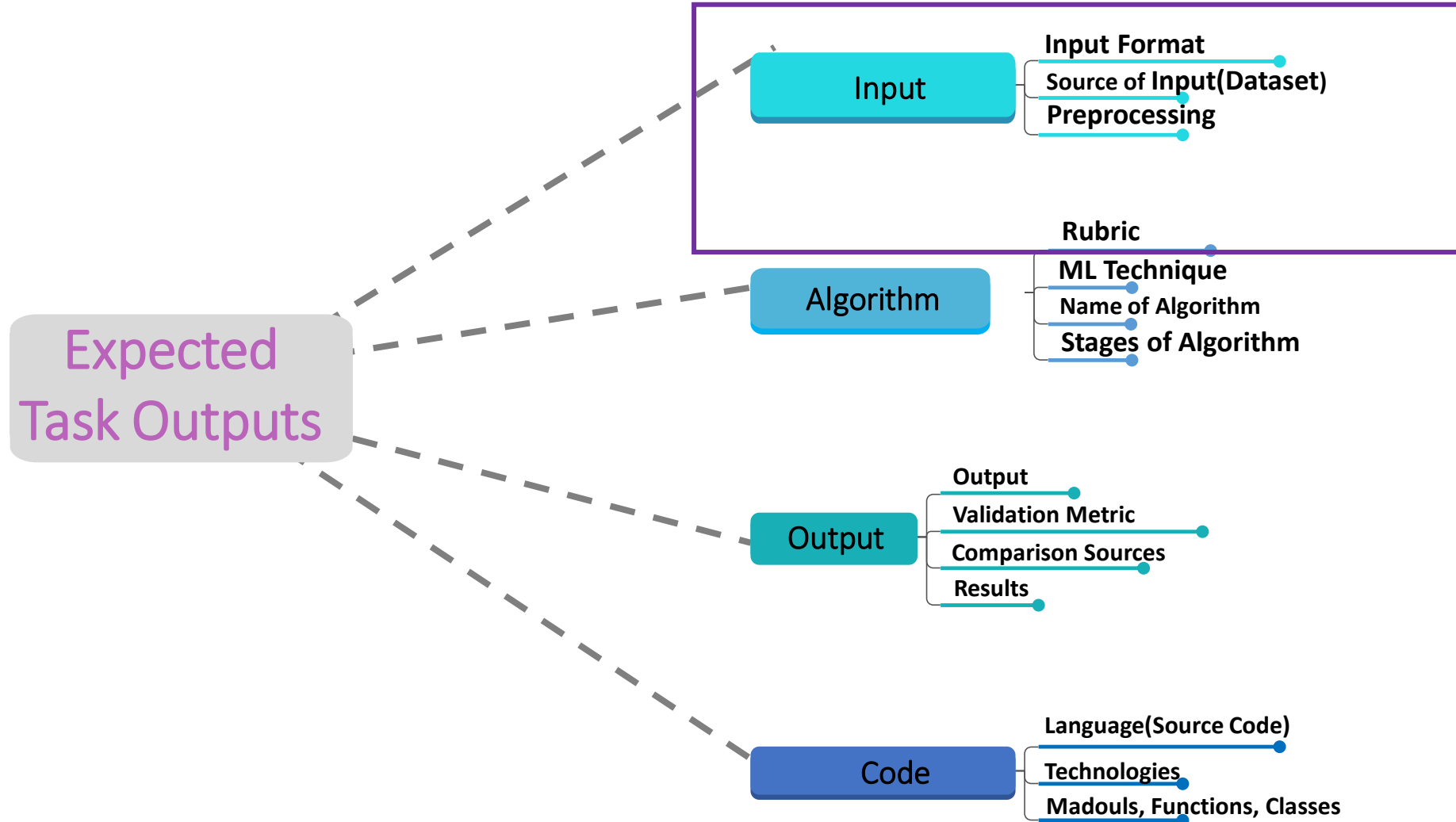


Task 5: Existing ML for Essay Scoring

Research and Code Analysis - Existing Work History In Three Different ML Approaches

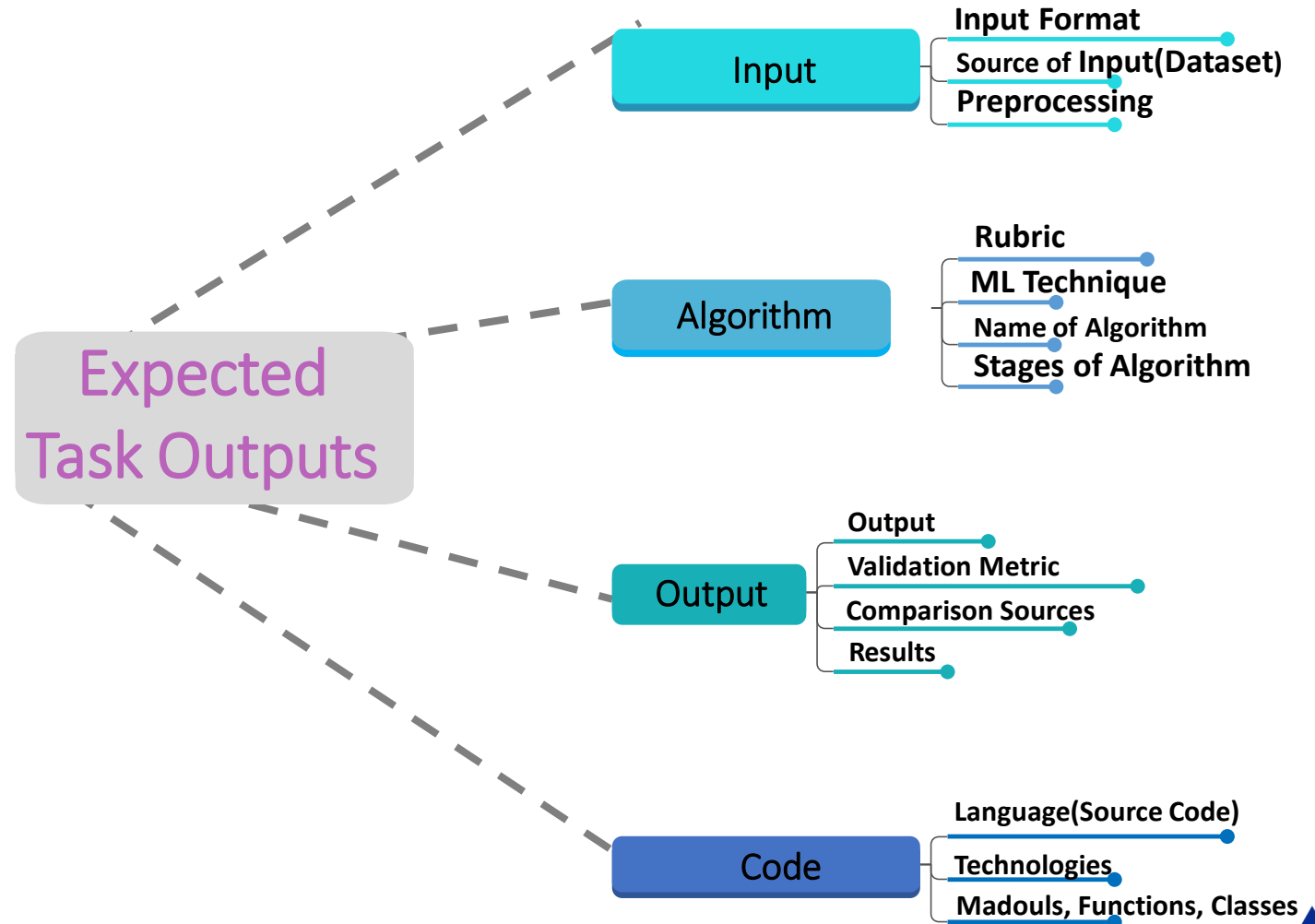
ML Approach	Name	Cited as	Source Code
Regression	EASE	Commercial	https://github.com/openedx-unsupported/ease
Neural Network	A neural approach to automated essay scoring.	Taghipour, K., & Ng, H. T. (2016, November). A neural approach to automated essay scoring. In Proceedings of the 2016 conference on empirical methods in natural language processing (pp. 1882-1891).	GitHub - nusnlp/nea: Neural Essay Assessor: An Automated Essay Scoring System Based on Deep Neural Networks
Neural Network	Automated scoring for reading comprehension via in-context bert tuning.	Fernandez, N., Ghosh, A., Liu, N., Wang, Z., Choffin, B., Baraniuk, R., & Lan, A. (2022, July). Automated scoring for reading comprehension via in-context bert tuning. In Artificial Intelligence in Education: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part I (pp. 691-697). Cham: Springer International Publishing.	https://github.com/ni9elf/automated-scoring
Classification	Effective feature integration for automated short answer scoring	Sakaguchi, K., Heilman, M., & Madnani, N. (2015). Effective feature integration for automated short answer scoring. In Proceedings of the 2015 conference of the North American Chapter of the association for computational linguistics: Human language technologies (pp. 1049-1054).	
	Automated Programming Assignment Marking Tool	2022 IEEE 7th International conference for Convergence in Technology (I2CT) Pune, India. Apr 07-09, 2022	

Task 5: Existing ML for Essay Scoring



Task 6: Research Bert Method

- <https://github.com/jina-ai/clip-as-service>



Task 7: Kaggle Analysis

9


Expected Task Output:

- Public Data Source Analysis
 - Exam
 - NLP Augmentation
- NLP Models and Codes
 - Available Model
 - Available Code

The image shows the Kaggle website header with navigation links: Competitions, Datasets, Models, Code, Discussions, Courses, and a search bar. Below this is a Jupyter Notebook titled "Predict Malicious Websites: XGBoost". The notebook code includes data loading, cleaning, splitting, and model fitting using XGBoost.

Start with more than a blinking cursor

Kaggle offers a no-setup, customizable, Jupyter Notebooks environment. Access GPUs at no cost to you and a huge repository of community published data & code.

 REGISTER WITH GOOGLE

Register with Email

```
data = pd.read_csv("../input/dataset.csv")

# clean up column names
data.columns = data.columns.\
    str.strip().\
    str.lower()

# remove non-numeric columns
data = data.select_dtypes(['number'])

# split data into training & testing
train, test = train_test_split(data, shuffle=True)

# peek @ dataframe
train.head()

# split training data into inputs & outputs
X = train.drop(["type"], axis=1)
Y = train["type"]

# specify model (xgboost defaults are generally fine)
model = xgb.XGBRegressor(tree_method = "gpu_exact")

# fit our model
model.fit(y=Y, X=X)

# split testing data into inputs & output
```