

HOW TO BUILD AN ML PORTFOLIO



QUT
AI & ML
SOCIETY

EVENTS OVERVIEW

WK 3

BUILDING AN ML PROFILE WITH **BANDS**

WK 4

MAINTAINING AN ML PROFILE WITH **BANDS**

WK 5

SPECIAL GENERAL MEETING

WK 6-8

PROJECT NIGHTS

WK 9

INDUSTRY SPEAKER



QUICK ACTIVITY



**STEP 1:
PULL OUT YOUR PHONE**



**STEP 2:
OPEN THE CAMERA APP**



SCAN ME



BECOME A MEMBER

Your membership helps us:

- Run workshops & events
- Host awesome networking nights
- Provide resources for our projects & activities



JOIN OUR TEAM

We now have the following vacancies

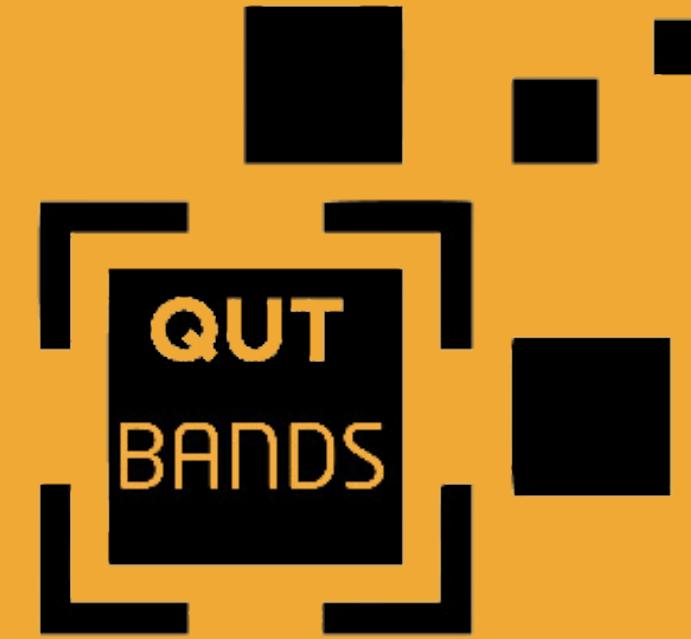
- Projects Officer (x3)
- Marketing Officer (Graphics) (x1)
- Events Officer (x1)
- General Executive (x1)



QUT BANDS

BUSINESS ANALYSIS & DATA SCIENCE CLUB

TRANSFORMING INSIGHTS INTO IMPACT



JOIN US

[LINKTR.EE/QUTBANDS](https://linktr.ee/qutbands)



@QUTBANDS





EVENT CALENDAR

Semester 2, 2025

Week 1. Amplify+ Hackathon

Together with a team, launch, build and pitch a solution to a real-world problem over the course of a weekend.

Week 2. Welcome Crafts Night

Get to know the community in a fun and casual night of crafting!

Week 3, 4. Projects Night w/ AMS

Dive deeper into your personal coding projects and collaborate with support from fellow students.

Week 5. Tech Industry Night 2025: Cyber, AI & You

Join us for an evening of insights, innovation, and industry connections that will shape the future of tech.

Week 6. BA Luncheon

Network with industry professionals over food as you explore the many ways Business Analysts can thrive.

Week 9. Speed Networking

Practice your interview skills in mock speed interviews. Meet with and pitch yourself to a variety of companies and get feedback!

Week 13. Movie Night w/ AMS and GEMS

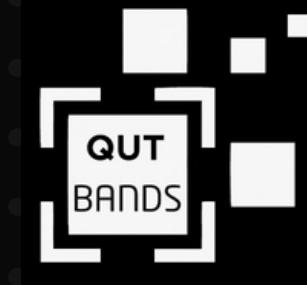
Kick back after a long semester with a night of movies and friends across the QUT community.

Weeks 2, 4, 6, 8 & 10. Coffee Catchup

Coffee catchups with various industry partners.

FOR TODAY

Right now, in all its chaotic glory



EVENT OVERVIEW

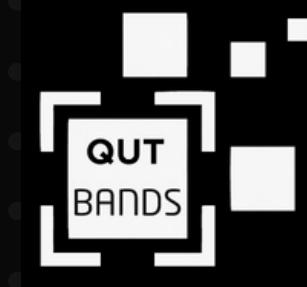
- Introduction to ML Portfolios
- GitHub for ML
- Where to find ML resources
- ML Project Documentation Best Practices
- ML Version Control and Collaboration with Git



With Jiya

INTRODUCTION TO ML PORTFOLIOS

The starter pack for ML greatness



WHAT IS A ML PORTFOLIO?

Machine Learning Portfolio

A curated collection of your machine learning projects showcasing your technical skills, problem-solving abilities, and creativity.

Recruiter Insight

Demonstrates to employers, recruiters, and collaborators your capabilities beyond coursework or job titles.

Practical Resume

Think of it as a hands-on resume – highlighting not just what you know, but what you've accomplished with your knowledge.



WHY DOES IT MATER?

Career Booster

Helps you stand out in internships, graduate programs, and job applications, even if you don't have much formal experience.

Show What You Can Do

Shows your drive, skills, and ability to explain your work clearly through project descriptions and notes.

From Idea to Finish

Proves you can take a project all the way through – finding data, building it, checking it, and sharing the results.



HOW TO FIND ML PROJECTS TO WORK ON AND CHALLENGES

Work on Real Problems

Focus on topics you care about – like sustainability, sports, finance, or healthcare.

Get Involved

Join hackathons, student clubs, or explore data on sites like Kaggle or the Australian Data Archive.

Start Simple

Begin by copying a well-known model, following a tutorial, or replicating a study. Then move on to your own ideas.



HOW TO FIND USEFUL RESOURCES FOR ML

Learn from Experts

Follow helpful YouTube channels like StatQuest, Data School, and MachineLearningStreetTalk, and take online courses on sites like Coursera or Fast.ai.

Stay Updated

Subscribe to newsletters such as The Batch or Towards Data Science to keep up with the latest news.

Explore Resources

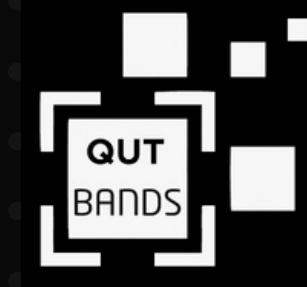
Use curated GitHub lists like awesome-machine-learning and read research papers on arXiv to discover new ideas and trends.



With Zac

GITHUB FOR ML

Branch out into machine learning



WHAT IS GITHUB

GitHub Platform

Store, track, and collaborate on code—your online project portfolio.

Recruiter Check

Review code quality, documentation, and project structure.

Team Coding

Supports version control, collaboration, and public visibility of your work.



HOW IS IT RELEVANT FOR ML

Model Hosting

Store and share machine learning models and experiments in one place.

Experiment Tracking

Keep version history of datasets, code, and model changes for reproducibility.

Community Sharing

Publish ML projects and collaborate with data scientists worldwide.



VIDEO: WHAT IS GITHUB



<https://www.youtube.com/watch?v=pBy1zgt0XPc>



CREATING A PROFESSIONAL GITHUB PROFILE

Profile Setup

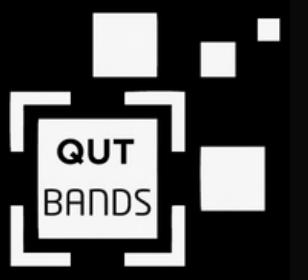
Use a clear photo, concise bio, and link to other professional profiles.

Pinned Projects

Highlight your best machine learning or data science work at the top.

Documentation & ReadMe

Add clean, well-structured READMEs to showcase clarity and professionalism.



DEMO: EXAMPLE PROFILES



jakevdp - Overview

Python ML & Data Science. jakevdp has 239 repositories available. Follow their code on GitHub.

 GitHub

[https://github.com/jakevdp?
tab=overview&from=2024-12-
01&to=2024-12-31](https://github.com/jakevdp?tab=overview&from=2024-12-01&to=2024-12-31)



codepo8 - Overview

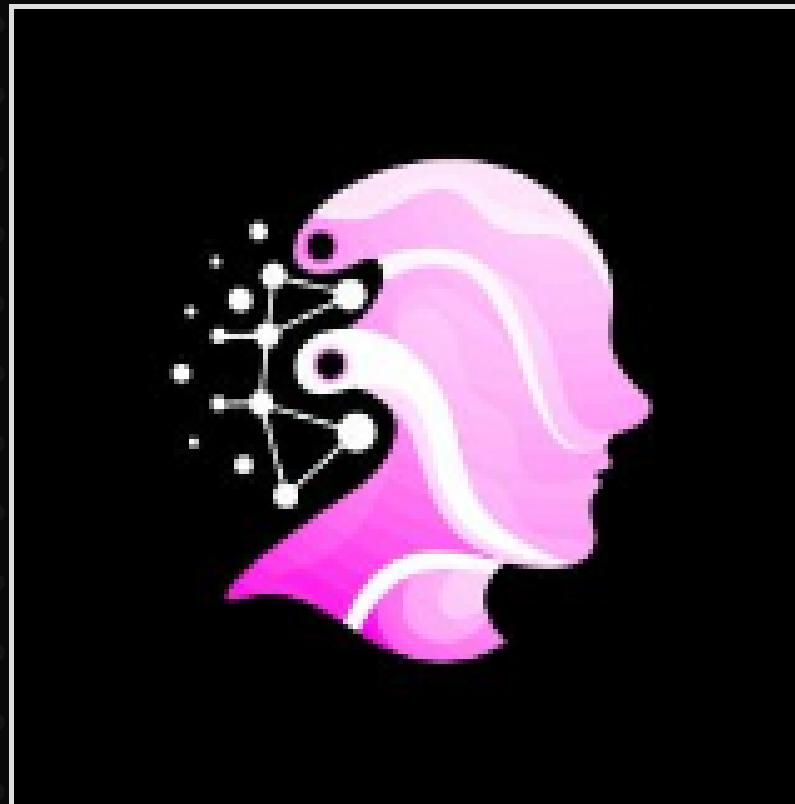
Engineer who worked for Microsoft/Mozilla/Yahoo on browsers, developer tools and anything JavaScript related. AI/ML enthusiast, writer and trainer. -....

 GitHub

<https://github.com/codepo8>



DEMO: AMS GITHUB



QUT AI & ML Society

Empowering the Future Through AI & Machine
Learning - QUT AI & ML Society



<https://github.com/QUT-AMS>



OPEN SOURCE CONTRIBUTION:

Why

Contributing to open source builds your credibility, helps you learn from senior developers, and gets your name out there.

How to Start

Pick beginner-friendly repos (look for "good first issue"), read their CONTRIBUTING.md file, and ask to be assigned a task.



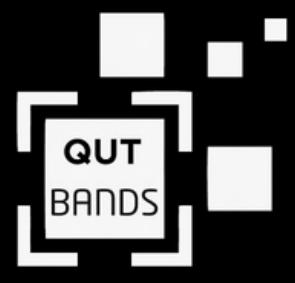
HOW TO STRUCTURE ML PROJECTS

- Organize **code**, **data**, and **notebooks** clearly for easy understanding and collaboration
- Include a **comprehensive README** to explain your project and usage
- Separate **raw data** from **processed data**; keep large files out of Git tracking



HOW TO STRUCTURE ML PROJECTS

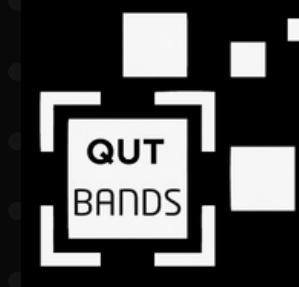
- Use **notebooks** for exploration and demos, **scripts** for production-ready code
- Commit often with **clear messages** and manage **versions** with tags
- **Modularize** code and **write tests** for reliability and reuse



With Jiya

WHERE TO FIND ML RESOURCES

Your treasure map to ML gold



WHAT IS KAGGLE?

A key platform for ML learners and professionals.

- Kaggle is a platform by Google for data science and machine learning competitions, datasets, notebooks, and community learning.
- Users can explore public datasets, collaborate on shared notebooks, and compete in real-world ML problems.
- Hot tip- Check out the official Kaggle youtube channel to get you started.



CREATING A STRONG KAGGLE PROFILE

- Bio: Include your background, interests (e.g. NLP, healthcare AI), and experience level.
- Notebooks: Share well-commented code with visualisations, markdown explanations, and proper structure.
- Competitions: Focus on learning, not just winning. Participate and document your thought process.
- Datasets: Clean datasets with good metadata get you visibility and medals.



COMPETITIONS: HOW TO START AND LEARN

Start with Getting Started competitions like:

- Titanic: <https://www.kaggle.com/c/titanic>
- House Prices: <https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

Use public notebooks to learn new techniques.

Work in teams for advanced comps.



NOTEBOOKS: SHARING YOUR ML WORK

- Use Kaggle Notebooks to:
- Document EDA, modeling, and evaluation
- Share insights via markdowns and charts
- Focus on storytelling, not just performance
- Style matters: clean code, clear headings, explanations

<https://www.kaggle.com/code/sisupalan/a-guide-to-write-crisp-and-concise-code>



DATASETS: UPLOADING AND USING THEM.

Content:

- Use datasets in your projects with citations
 - Create your own dataset from public APIs, web scraping, or CSVs
 - Well-documented datasets can earn you medals and followers
-
- Kaggle Datasets Page: <https://www.kaggle.com/datasets>



DISCUSSION FORUMS AND LEARNING COMMUNITY,

- Ask questions, join threads, and engage with experts
- Learn best practices and discuss competition strategies
- Follow Grandmasters to track their activity and get inspired



HOW TO LINK KAGGLE ACHIEVEMENTS

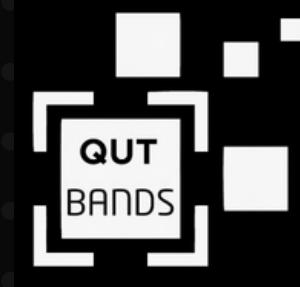
- Add competition rankings, medal wins, and dataset contributions to your resume.
- Link your Kaggle profile on LinkedIn and GitHub.
- Screenshot leaderboard achievements if they're notable.



With Zac

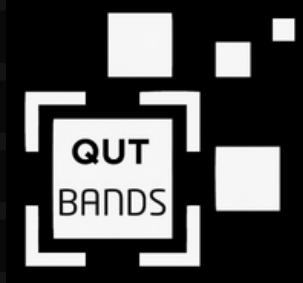
ML PROJECT DOCUMENTATION FOR BEST PRACTICES

Docs that don't make you cry



WHY STRUCTURE YOUR ML PROJECTS

- **Clear communication** - Helps others understand your work easily.
- **Professionalism** - Shows you know best practices for coding and project management.
- **Reproducibility** - Makes it simple to rerun experiments and validate results.
- **Collaboration** - Enables others to contribute without confusion.



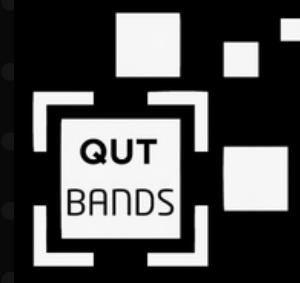
TYPICAL ML PROJECT STRUCTURE

```
project-name/
|
+-- data/
|   +-- raw/          # Original unprocessed data
|   +-- processed/    # Cleaned and transformed data
|
+-- notebooks/       # Jupyter notebooks for exploration and demos
|
+-- src/
|   +-- data/         # Data loading and preprocessing scripts
|   +-- features/     # Feature engineering code
|   +-- models/        # Model training and evaluation code
|   +-- utils/         # Utility functions
|
+-- tests/           # Unit tests and integration tests
|
+-- requirements.txt  # Python dependencies
+-- README.md         # Project overview and instructions
+-- LICENSE           # License info
+-- setup.py          # Installation setup (optional)
+-- .gitignore         # Files to exclude from Git tracking
```



README ESSENTIALS

- Project description and goals
- Setup instructions (install dependencies, data download)
- How to run code/notebooks
- Explanation of folder structure
- Results summary and visualizations
- Contribution guidelines



EXAMPLE README

QUT-AMS/Linear-Regression



This repository is a beginner-friendly introduction to linear regression, designed to help users understand one of the foundational algorithms in...

2 Contributors 0 Issues 0 Stars 0 Forks

<https://github.com/QUT-AMS/Linear-Regression>

QUT-AMS/Linear-Regression: This repository is a beginner-friendly introduction to linear regression,...

This repository is a beginner-friendly introduction to linear regression, designed to help users understand one of the foundational algorithms in machine learning. - QUT-AMS/Linear-Regression

 GitHub

<https://github.com/QUT-AMS/Linear-Regression>



DATASET DESCRIPTION

- **Source** - Where the data comes from (e.g., public dataset, company data, web scraping)
- **Content** - What the dataset contains – features, labels, sample size, and types of data
- **Format** - File types (CSV, JSON, images, etc.) and folder structure for datasets
- **Preprocessing** - Summary of cleaning, filtering, and transformation steps applied
- **Usage Instructions** - How to load and use the dataset within your project (code examples or scripts)



MODEL EXPLANATION

- **Model Type:** Describe the algorithm(s) used (e.g., Random Forest, CNN, Transformer)
- **Architecture:** Outline model architecture, layers, and key parameters (especially for deep learning)
- **Training Details:** Dataset splits, hyperparameters, training duration, and hardware used
- **Performance Metrics:** Key evaluation metrics (accuracy, F1-score, RMSE, etc.) and validation methods



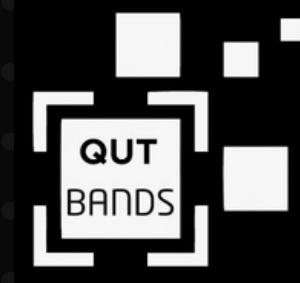
RESULTS + VISUALS

- **Summary of Outcomes:** Clear and concise description of model performance and key findings
- **Graphs & Charts:** Include plots like confusion matrices, ROC curves, training/validation loss, feature importance
- **Tables:** Present numeric results such as metric scores, hyperparameter comparisons, or experiment logs
- **Visualizations:** Use images, sample predictions, or example outputs to demonstrate model behavior
- **Interpretation:** Explain what the results mean in the context of your problem and any next steps



USE NOTEBOOKS EFFECTIVELY

- Use notebooks for exploration, visualisation, and demoing results
- Keep notebooks clean and linear (avoid long messy outputs)
- Export key code from notebooks to scripts in src/ for production use
- Include comments and markdown explanations



NOTEBOOKS EXAMPLE

salma2vec/ML-Beginner-Portfolio

Kickstart ML through these 20+ foundational projects; Kaggle datasets, problem statements and comprehensive EDA (Exploratory Data Analysis) walkthroughs.

Ax 1 ⏺ 0 ★ 8 ⚡ 1
Contributor Issues Stars Fork



salma2vec/ML-Beginner-Portfolio: Kickstart ML through these 20+ foundational projects; Kaggle datasets,...

Kickstart ML through these 20+ foundational projects; Kaggle datasets, problem statements and comprehensive EDA (Exploratory Data Analysis) walkthroughs. - salma2vec/ML-Beginner-Portfolio

 GitHub

<https://github.com/salma2vec/ML-Beginner-Portfolio>



BEST PRACTICES FOR CODE

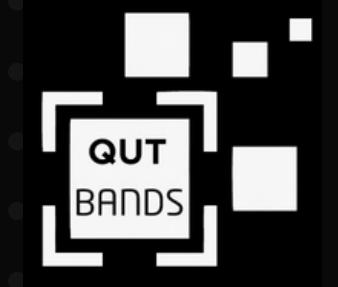
- Keep raw data separate and never commit large raw files
- Use scripts for preprocessing instead of manual notebook steps
- Modularise code for reusability
- Write tests to validate key functions
- Use meaningful commit messages and document changes



Also with Zac

VERSION CONTROL AND COLLABORATION WITH GIT

Keep your code together (literally)



DEMO: GITHUB APP



COLLABORATION AND OPEN SOURCE CONTRIBUTION

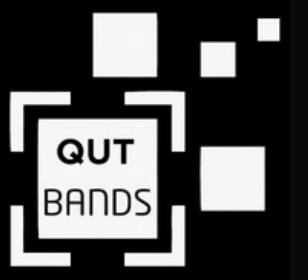
Why contribute?

Builds credibility, improves skills, and expands your network

How to start?

Find repos labeled “good first issue”, read contributing guidelines, ask for tasks

Use Pull Requests with clear descriptions and reference issues



WHY USE VERSION CONTROL IN ML PROJECTS

- Track changes to code, data preprocessing, and experiments
- Collaborate smoothly with team members and open source contributors
- Easily revert to previous versions if something breaks
- Manage different branches for feature development and experimentation
- Maintain reproducibility of models and results



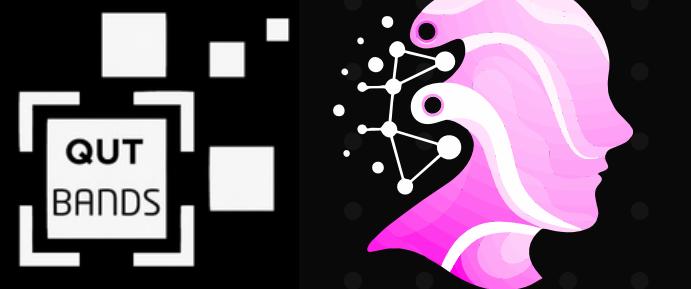
WHAT TO TRACK AND WHAT TO IGNORE

Track

- Code
- Experiment configs
- Small sample datasets
- Training scripts
- Evaluation metrics

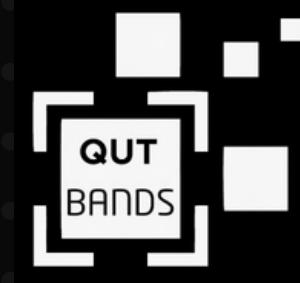
Ignore

- Large raw datasets
- Model checkpoints
- Logs
- Temporary files



WRITING EFFECTIVE COMMIT MESSAGE

- Be clear and descriptive about what changed and why
- Use present tense verbs (e.g., “Add data preprocessing script”)
- Reference related issues or tasks when possible
- Commit often with focused, small changes for easier reviews



BRANCHING AND MERGING

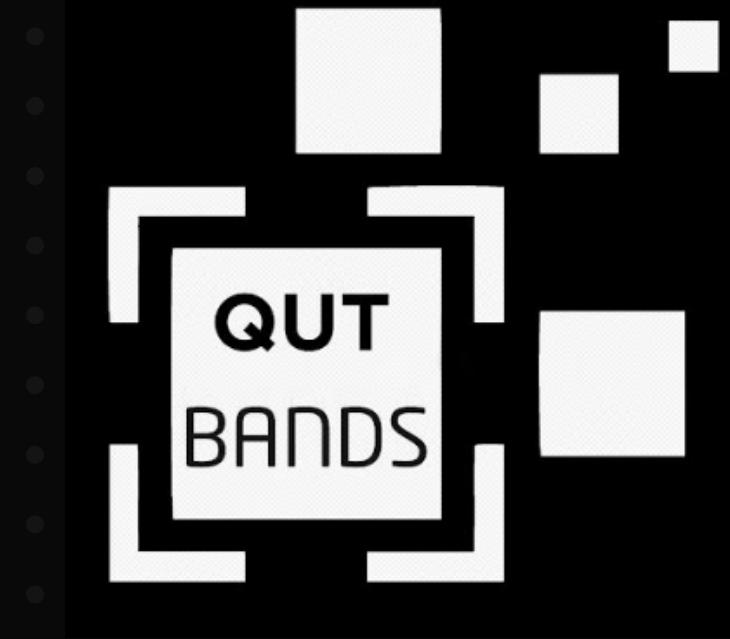
- Use branches for new features, experiments, or bug fixes
- Keep the main branch stable and deployable
- Merge frequently to avoid complex conflicts
- Use Pull Requests to review code before merging



BRANCHING AND MERGING

- Use branches for new features, experiments, or bug fixes
- Keep the main branch stable and deployable
- Merge frequently to avoid complex conflicts
- Use Pull Requests to review code before merging

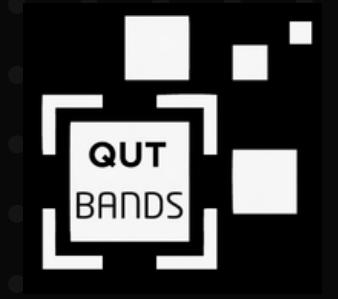




THANK YOU FOR COMING

WHAT'S NEXT

Upcoming Events



UPCOMING EVENT



The poster features a black background with a subtle dotted pattern. At the top, there's a yellow diagonal bar. In the center, there are two logos: a pink profile of a head with a brain-like network inside, and a white square icon with the text "QUT BANDS" and a geometric pattern. Below the logos, the words "PROJECTS NIGHT" are written in large, bold, yellow capital letters. Underneath that, in a smaller orange rounded rectangle, is the text "How to Maintain a ML Portfolio 101". At the bottom left, there's event information: a calendar icon followed by "Tue 12/8, 5:30PM-7:30PM", and a location pin icon followed by "QUT GP-B223". At the very bottom, the text "BUILD YOUR ML PORTFOLIO WITH PROJECT NIGHTS: LEARN. CREATE. SHOWCASE." is displayed in white.



UPCOMING EVENT

No registration required



Coffee catchups



with
Deloitte.

MONDAY 10-11AM
Week 4

The Kitchen, QUT Gardens Point



2025 QUT TECH INDUSTRY NIGHT

CYBER, AI & YOU

Dive into the frontlines of tech where digital defence meets intelligent design, and discover how you can shape a safer tomorrow.



GET YOUR
TICKETS HERE!

FRIDAY 22nd August

5:30 PM - 10 PM

**Gardens Theatre,
X Block QUT Gardens Point**