Supervisor	Start Date	End Date	Student	Main Goal	Why I Chose This Project
Mohammed Amin Adoul	2025-09-15	2025-12-12	Studenka Lundahl	Learn EEG signal processing, extract features, apply ML/DL classifiers (logistic regression, Random Forest, CNN/LSTM), and evaluate performance in predicting 'Hit' vs 'Miss' outcomes.	I am taking several courses this autumn and also preparing for knee surgery. I chose EEG because it is well structured with clear milestones, focused on ML/DL methods, and manageable alongside my other commitments.

Project title: Signal processing and machine learning for EEG data classification

Task Name	Status	Duration [days]	Start Date	End Date	Tools Planned	Comments
PHASE 1: FOUNDATION & GRADE 3						
1.1 Project Decision & Project plan creation	100%	9	2025-09-03	**2025-09-12**	MS Excel	
1.2 Recovery & theory study (surgery week)	100%	4	2025-09-13	2025-09-16	Reading materials, MNEdocs	
1.3 Install libraries & environment setup	100%	2	2025-09-17	2025-09-18	Python, MNE, pandas, matplotlib	
1.4 Data exploration & understanding	100%	3	2025-09-19	2025-09-21	MNE, pandas, matplotlib	
1.5 Basic data preprocessing (filtering, artifacts)	100%	4	2025-09-22	2025-09-25	MNE, NumPy, SciPy	
1.6 Event alignment (Hit/Miss mapping)	100%	2	2025-09-26	2025-09-27	MNE, pandas	
1.7 Basic feature extraction	100%	2	2025-09-28	2025-09-29	MNE, NumPy	
1.8 Logistic regression implementation	100%	2	2025-09-30	2025-10-01	scikit-learn	
Milestone 1 - Report & Demo (Grade 3)	100%	1	**2025-10-01**	**2025-10-01**	–	
PHASE 2: ADVANCED & GRADE4						
2.1 Advanced feature extraction (PSD)	100%	4	2025-10-02	2025-10-05	MNE, SciPy	
2.2 Spectral entropy & wavelet features	0%	4	2025-10-06	2025-10-09	SciPy, PyWavelets	Status: 0% (Intentionally skipped) 152 PSD features provided sufficient discriminative power. Additional features would risk overfitting at current sample size.

2.3 Connectivity features between channels	0%	3	2025-10-10	2025-10-12	MNE, NumPy	Status: 0% (Intentionally skipped) 152 PSD features provided sufficient discriminative power. Additional features would risk overfitting at current sample size.
2.4 Random Forest implementation	100%	2	2025-10-13	2025-10-14	scikit-learn	
2.5 Feature importance analysis	100%	2	2025-10-15	2025-10-16	scikit-learn, matplotlib	
2.6 Model comparison (LR vs RF)	100%	1	2025-10-17	2025-10-17	scikit-learn	
Milestone 2 - Report & Demo (Grade 4)	100%	1	**2025-10-17**	**2025-10-17**	–	
PHASE 3: DEEP LEARNING PREP						
3.1 Data preparation for deeplearning	100%	3	2025-10-18	2025-10-20	TensorFlow/Keras, NumPy	
3.2 CNN model development	100%	4	2025-10-21	2025-10-24	TensorFlow/Keras	
3.3 CNN training & validation	100%	3	2025-10-25	2025-10-27	TensorFlow/Keras	
3.4 LSTM model development	0%	3	2025-10-28	2025-10-30	TensorFlow/Keras	Supervisor feedback: LSTM not necessary since it's almost like CNN. Time invested in comprehensive three-model statistical analysis instead, providing greater scientific value.
Milestone 3 - Report & Demo (Basic DL Grade 5)	100%	1	**2025-10-30**	**2025-10-30**	–	Per supervisor guidance, LSTM not required (similar to CNN). Focus shifted to comprehensive three-model statistical comparison and report quality. CNN fully implemented (61% accuracy). Grade 5 deep learning requirements completed.
PHASE 4: ADVANCED DEEP LEARNING						
4.1 LSTM optimization & validation	0%	5	2025-10-31	2025-11-04	TensorFlow/Keras	
4.2 Transformer model (optional)	0%	7	2025-11-05	2025-11-11	TensorFlow/Keras	
4.3 Hyperparameter optimization	50%	5	2025-11-12	2025-11-16	TensorFlow/Keras, Optuna	Completed: RF hyperparameters, CNN early stopping, learning rate scheduling Not completed: Systematic grid search (not needed at current data scale)
4.4 Final model comparison & analysis	100%	3	2025-11-17	2025-11-19	All tools	

4.5 Performance optimization	0%	2	2025-11-20	2025-11-21	All tools	
Milestone 4 - Report & Demo (Grade 5)	100%	1	**2025-11-21**	**2025-11-21**	–	Supervisor-approved scope focused on comprehensive analysis and report quality. Grade 5 exceeded. Three-model statistical comparison (LR vs RF vs CNN) completed with effect size analysis, environmental factors investigation, and publication-quality evaluation.
FINAL PHASE						
5.1 Final report writing (IMRAD format)	100%	15	2025-11-22	2025-12-06	MS Word, matplotlib	
5.2 Code documentation & cleanup	100%	3	2025-12-07	2025-12-09	Jupyter, GitHub	
5.3 Final presentation preparation	100%	2	2025-12-10	2025-12-11	PowerPoint, matplotlib	
Final - Report & Demo	100%	1	**2025-12-12**	**2025-12-12**	–	

Status
0% = Not started
25% = Started
50% = In progress
75% = Nearly complete
100% = Complete

Risk	Mitigation Strategy
Knee surgery recovery (Sept 11-22)	Plan light workload with reading/theory only. No hands-on coding during recovery week.
Time overlap with other courses	EEG chosen for clear milestones and manageable workload. Focus on Grade 3 first as minimum viable outcome.
Deep learning models too complex/time-consuming	If time runs out, focus on ML methods (logistic regression, Random Forest) for Grade 4. DL is bonus for Grade 5. RESOLVED: Focused on CNN implementation with supervisor-approved scope. LSTM/Transformer deprioritized per academic guidance in favor of rigorous statistical comparison.
EEG data preprocessing challenges	Start simple, use MNE tutorials extensively, allocate extra debugging time in early phases.
Computational requirements for deep learning	Use Google Colab for GPU access, start with smaller data subsets, have simpler models as backup.
Multiple course deadlines conflicting	Prioritize Grade 3 completion by end October, treat Grade 4-5 as stretch goals if time permits.
Technical difficulties with MNE library	Join MNE community forums, use extensive documentation, plan buffer time for learning curve.
Poor model performance on Hit/Miss classification	Try different preprocessing approaches, feature engineering, and ensure proper cross-validation setup.

==========
ACHIEVEMENTS BEYOND ORIGINAL PLAN
==========

- ☑ Three-model statistical comparison (LR vs RF vs CNN)
- ☑ Effect size analysis (Cohen's d): RF vs LR d=-0.57 (medium effect)
- ☑ Environmental factors: r=0.605, p=0.064 correlation identified
- ☑ Comprehensive feature importance: 152 PSD features analyzed
- ☑ Publication-quality report: 6-page academic paper (IMRAD format)
- ☑ Professional presentation: Final project presention - 21-minute talk

Strategic Decisions:
- LSTM deprioritized → Three-model comparison (better scientific value)
- Connectivity features skipped → Focused on PSD depth (152 features)
- Systematic grid search → Manual tuning adequate for data scale

Result: Project demonstrates graduate-level research quality with appropriate statistical rigor. All Grade 3, 4, and 5 requirements completed and exceeded.