

CSCI 490 Senior Research Project Outline

Student: Chris Lawrence

Project Title: Astronomical Identification: Temporal Dataset Preparation

Supervisor: Prof. Luis Meneses

Co-Supervisor: [CSCI Faculty Member TBD]

Committee members: Prof. Gregory Arkos

Proposal: CSCI 490 Senior Research Project (3 credits)

Duration: September 2025 – December 2026

Topic Outline

The AstrID project was partially developed as a prototype for identifying stars in astronomical images using FITS data and a U-Net machine learning model, developed for a CSCI 478 course in F24 semester. This existing system can successfully localize and classify stars with reasonable accuracy, serving as a foundation for further research.

This project will extend AstrID in one key area:

Temporal Dataset Preparation – developing a comprehensive pipeline for processing time-series astronomical observations, cleaning and preparing temporal datasets, to prepare for astronomical identification of events (e.g., supernovae, brightness variations) across multiple observations.

This builds directly on the completed AstrID framework while focusing on the critical foundation work needed for temporal analysis.

Project Description

The project will allow me to learn and utilize a systematic approach combining data engineering, machine learning research, and astronomical data analysis to establish a robust temporal dataset pipeline. A course that combines these learning outcomes is not offered at VIU, which is why I am proposing them as a Senior Research Project. Also, the skills that I will learn are very sought after in today's industry.

Core Research Components:

- **Temporal Dataset Preparation:** Investigation and implementation of data cleaning, preprocessing, and ingestion techniques for time-series astronomical imaging data. This includes developing robust pipelines for handling FITS format images across multiple observations, ensuring data quality, and creating standardized temporal datasets.

Key Implementation Areas:

- Temporal data ingestion and validation pipelines
- Data quality assessment and cleaning algorithms
- Time-series alignment and normalization techniques

Research Methodology: The project will employ a focused approach combining:

- Literature review of astronomical time-series data processing
- Development and testing of temporal data preparation pipelines
- Quantitative evaluation with data quality metrics
- Iterative refinement based on pipeline performance results

Evaluation Framework: The research will be evaluated through multiple criteria including:

- Technical functionality and reliability of temporal data processing pipeline
- Data quality metrics and consistency across time-series observations
- Scalability and generalizability of the temporal solution to other astronomical datasets

Project Criteria for Different Grade Levels:

- **A+ (Pass with honours):** Exceptional temporal data pipeline reliability, comprehensive documentation, and clear contribution to temporal astronomical data processing methodology.
- **B+ (Pass with distinction):** High temporal data pipeline reliability, thorough documentation, and some contribution to temporal data processing.
- **C (Pass):** Working temporal data pipeline, adequate documentation, and demonstrated learning outcomes.
- **F (Fail):** Incomplete or non-functional temporal pipeline, inadequate documentation, or failure to meet minimum project requirements.

Future Research Directions: The temporal dataset framework established in this project will enable future investigations into:

- Advanced anomaly detection algorithms for astronomical time-series
- Real-time transient event detection in astronomical surveys

Project Deliverables and Schedule

The student will meet weekly with Prof. Meneses (schedule TBA) to review:

- Progress on milestone goals
- Technical challenges or architectural decisions
- Feedback on code, documentation, and presentation preparation

At each meeting, the student will provide a short (10-15 minute) briefing on the project state and bring current copies of relevant documents.

Fall 2025 Project Schedule:

Date	Deliverable
Sept16	Finalize proposal and registration form; project kickoff; environment setup
Sept	Literature review on temporal astronomical data processing ; initial research log
Oct	Dataset preparation and cleaning pipeline development; temporal data validation framework

Date	Deliverable
Nov	Implement temporal data ingestion system; preliminary evaluation
Dec	Complete temporal pipeline; final evaluation and testing; progress presentation

Final Deliverables

- **Complete AstrID Temporal Dataset Pipeline System**
- **Final Project Report documenting:**
 - Astronomical temporal data processing problem domain and existing solution analysis
 - Research methodology and data engineering approach
 - Technical implementation of temporal data pipelines algorithms
 - Performance evaluation results and data quality metrics
 - Limitations and future expansion opportunities
 - Critical analysis of approach and methodology
- **Oral Defense and demonstration of temporal data processing capabilities**

Additional Notes

- The entire project will be hosted and versioned via GitHub
- All findings must be documented in a detailed, carefully written and reviewed project report
- The final project report will include both formal description of the project, the research behind it, the practical implementation, rationale behind implementation choices, and experimentation showing strengths and weaknesses of chosen approaches
- This project serves as a foundation for CSCI 491 continuation in the following semester