October 14, 2023

C964: Computer Science Capstone Template

**Note:** This is the latest version of the Task 2 template. Following this template meets all the documentation requirements for C964 version SIM2 and SIM3. As it’s more succinct and clear, we recommend using this template for both SIM2 and SIM3. However, using the [previous template](https://westerngovernorsuniversity-my.sharepoint.com/:w:/g/personal/jim_ashe_wgu_edu/EcklZjLXTB5EpDS4BVYc8SEBhT3VHy3s_9lZSIZ5aH6Q5w?e=5tCTQb) is still acceptable.

Task 2 parts A, B, C and D

[Part A: Letter of Transmittal 1](#_Toc98598250)

[Letter of Transmittal Requirements 2](#_Toc1738085866)

[Letter Template 2](#_Toc1133266111)

[Part B: Project Proposal Plan 3](#_Toc1370766476)

[Project Summary 4](#_Toc904507251)

[Data Summary 4](#_Toc1736393957)

[Implementation 4](#_Toc1241988654)

[Timeline 4](#_Toc1357178365)

[Evaluation Plan 5](#_Toc623361460)

[Resources and Costs 5](#_Toc1538507987)

[Part C: Application 5](#_Toc1471073175)

[Part D: Post-implementation Report 6](#_Toc651895932)

[Solution Summary 7](#_Toc1134136520)

[Data Summary 7](#_Toc182221765)

[Machine Learning 7](#_Toc1505466430)

[Validation 7](#_Toc391434166)

[Visualizations 7](#_Toc201059345)

[User Guide 7](#_Toc1365484010)

[Reference Page 8](#_Toc1702353417)

# Part A: Letter of Transmittal

## Letter of Transmittal Requirements

The *Letter of Transmittal* should convince senior leadership to approve your project. Write a brief cover letter (suggested length 1-2 pages) describing the problem, how the application (part C) applies to the problem, the practical benefits to the organization, and a brief implementation plan. Include all artifacts typical of a professional (business) letter, e.g., subject line, date, greeting, signature, etc.

The letter should be concise and target a non-technical audience. Include the following:

**A summary of the problem.**

At the heart of education lies a shared objective among educators: to unlock every student's academic potential. This goal demands proactive strategies to uplift students across the academic spectrum—from high achievers to those requiring additional support. However, Valley High School (VHS), our client, has historically operated reactively due to the absence of predictive tools, often reacting only after student performance is evident. Adding to the challenge is the holistic nature of student academic performance, which is influenced by both academic and non-academic factors. Yet, VHS’s current analytical tools overlook this comprehensive perspective, focusing solely on academic lag measurements such as test scores and absences. Recognizing our client's objective and current toolset the problem is clear: VHS lacks a holistic predictive application necessary to proactively support its students.

## Letter Template

May 17, 2024

[Recipient’s name]

MindShift Solution Experts

839 Main Street

Santa Ana, CA 92705

Dear [Recipient’s name],

At the heart of education lies a shared objective among educators: to unlock every student's academic potential. This goal demands proactive strategies to uplift students across the academic spectrum—from high achievers to those requiring additional support. However, Valley High School (VHS), our client, has historically operated reactively due to the absence of predictive tools, often reacting only after student performance is evident. Adding to the challenge is the holistic nature of student academic performance, which is influenced by both academic and non-academic factors. Yet, VHS’s current analytical tools overlook this perspective, focusing solely on academic lag measurements such as test scores and absences. Recognizing our client's objective and current toolset the problem is clear: VHS lacks a holistic predictive application necessary to proactively support its students.

Praesent sagittis, leo vitae sodales cursus, purus orci rhoncus ligula, quis facilisis ante est nec lectus. Curabitur augue quam, ultricies at arcu eget, molestie eleifend neque. Sed interdum tempor purus, et luctus tellus feugiat imperdiet. Maecenas scelerisque viverra orci tincidunt luctus. Vestibulum sodales eros ut ex luctus tempor. Curabitur eget leo vehicula, malesuada urna ut, eleifend nisi. Quisque sapien tellus, ornare ac magna quis, ultricies consectetur mauris. Nunc erat ligula, mattis id tempor ut, venenatis ac mi. Sed sit amet odio ac ligula tincidunt iaculis sit amet vel ipsum.

Etiam lobortis aliquam metus, eu aliquam ante aliquam ut. Nam tristique sagittis mauris vel tempor. Quisque rhoncus, justo sed lobortis porta, nulla libero pulvinar tortor, ut ullamcorper justo enim ut erat. Praesent lobortis ut leo in aliquet. Suspendisse aliquet velit nulla, a rhoncus nibh vestibulum iaculis. Praesent mollis nibh nec ultrices blandit. Pellentesque felis elit, pretium at risus in, commodo consectetur tortor. Etiam fringilla mi quis erat mollis ultricies. Phasellus vestibulum elementum commodo. Sed congue vulputate orci in porta. Pellentesque scelerisque facilisis justo, a bibendum ligula tempus quis. Aenean efficitur eleifend lorem, et tempus risus consequat quis. Cras varius metus sapien, ac malesuada sem volutpat sed.

Sincerely,

[Sign here: e.g., Jane Smith]

[Your name, title]

# Part B: Project Proposal Plan

## Project Summary

Our client, Valley High School (VHS) currently relies on tools and methods that focus on lag indicators, such as past grades and absences, to improve academic performance. This has lead to an overall strategy that is reactive in nature. However, research has established that proactive academic strategies are critical for enhancing student performance. To implement proactive strategies, the school must be able to predict student performance. Recognizing the limitations of the current tools available to VHS, the problem is evident: VHS cannot implement proactive academic improvement strategies effectively due to the lack of predictive capabilities.

To address this, we propose the development of an application that utilizes data analysis and machine learning to forecast academic performance based on a range of factors. The application will analyze and train on a mix of data, academic (e.g., test scores, attendance) and non-academic (e.g., health, social interactions). The tool will then be able to predict the student’s test score given selected attributes.

To meet VHS’s needs, we will develop and provide the following deliverables:

* **Jupyter Notebook Application:** A detailed, interactive notebook containing all necessarycode, data analyses, and machine learning algorithms.
* **User Guide:** Comprehensive instructions for running the Jupyter Notebook application,including steps for conducting analyses and interpreting results.

Project Summary **(continued)**

This application will provide VHS with the foresight gained from machine learning and predictive analytics, enabling the school to proactively implement academic-enhancing measures. The implementation of this application is expected to significantly improve student outcomes, reduce failures, and enhance overall educational effectiveness at VHS for all students, regardless of academic performance.

To address this, we propose developing an application that utilizes data analysis and machine learning to predict future individual academic performance. This tool will analyze a mix of academic (e.g., test scores, attendance) and non-academic factors (e.g., health, social interactions), providing a comprehensive approach.

To meet VHS’s needs, we will develop and provide the following deliverables:

* **Jupyter Notebook Application:** A detailed, interactive notebook containing all necessarycode, data analyses, and machine learning algorithms.
* **User Guide:** Comprehensive instructions for running the Jupyter Notebook application,including steps for conducting analyses and interpreting results.

## 

This application will provide VHS with the foresight gained from machine learning and predictive analytics, enabling the school to proactively implement academic-enhancing measures. The implementation of this application is expected to significantly improve student outcomes, reduce failures, and enhance overall educational effectiveness at VHS for all students, regardless of academic performance.

## Data Summary

To ensure the project's proper alignment with the objectives, the raw data will be sourced from an open-source dataset already obtained from Kaggle.com. This particular dataset was selected for its robustness and the comprehensive variety of data collected, which encompasses both academic and non-academic factors crucial for our predictive analysis. Moreover, the dataset has been preprocessed to remove any personally identifying information, ensuring compliance with privacy regulations.

Next Paragraph:

Describe how data will be processed and managed throughout the application development life cycle: design, development, maintenance, or others.

Justify why the data meets the needs of the project. If relevant, describe how data anomalies, e.g., outliers, incomplete data, etc., will be handled.

Address any ethical or legal concerns regarding the data. If there are no concerns, explain why.

## Implementation

* Describe an industry-standard methodology to be used.
* An outline of the project’s implementation plan. This outline can focus on the project’s development as a whole; or it may focus on only the implementation of the machine learning solution.

## Timeline

* Provide a projected timeline, including projected start dates and end dates for each milestone (a table is not required but encouraged).

|  |  |  |  |
| --- | --- | --- | --- |
| Milestone or deliverable | Duration  (hours or days) | Projected start date | Anticipated end date |
|  |  |  |  |
|  |  |  |  |

## Evaluation Plan

* Describe the verification method(s) to be used at each stage of development.
* Describe the validation method to be used upon completion of the project.

## Resources and Costs

* Itemize hardware and software costs.
* Itemize estimated labor time and costs.
* Itemize estimated environment costs of the application, e.g., deployment, hosting, maintenance, etc.

# Part C: Application

Part C is your submitted application. This part of the document can be left blank or used to include a list of any submitted files or links.

The minimal requirments of the submittred *application* are as follows:

1. **The applicaton functions as described.** Following the ‘User Guide’ in part D, the evaluator must be able to succesfully review your application on a Windows 10 machine.
2. **A mathematical algorithm applied to data,** e.g, supervised, unsupervised, or reinforced machine learning method.
3. **A “user interface.”** Following the ‘User Guide’ in part D, the client must be able to use the application towards solving the proposed problem (as described in parts A, B, and D). For example, the client can input varaibles and the application oututs a prediction.
4. **Three visualizations.** The visualizations can be included separately when including them in the application is not ideal or possible, e.g., the visualizations describe proprietary data but the application is customer-facing.
5. **Submitted files and links are static and accessible.** All data, source code, and links must be accessible to evaluators on a Windows 10 machine. If parts of the project are able to be modified after submission, then matching source files must be submitted. For example, if the application is a website or hosted notebook, the .html or .ipynb files must be submitted directly to assessments.

Ideally, submitted applications should be reviewable using either Windows or Mac OS, e.g., Jupyter notebooks, webpages, Python projects, etc. If the source files exceed the 200 MB limit, consider providing screenshots or a Panopto video of the functioning application and contact your course instructor.

# Part D: Post-implementation Report

Create a post-implementation as outlined below. Provide sufficient detail so that a reader knowledgeable in computer science but unfamiliar with your project can understand what you have accomplished. Using examples and visualizations (including screenshots) beyond the three required is recommended (but not required). **Write everything in the past tense.**

## Solution Summary

* Summarize the problem and solution.
* Describe how the application provides a solution to the problem from parts A and B.

## Data Summary

* Provide the source of the raw data, how the data was collected, or how it was simulated.
* Describe how data was processed and managed throughout the application development life cycle: design, development, maintenance, or others.

## Machine Learning

For each employed method (at least one is required) provide the following:

* Indentify the method and what it does (the “what”).
* Descirbe how the method was developed (the “how”).
* Justifiy the selection and developement of the method (the “why”).

## Validation

For each employed method described in the section above provide the following:

* An appropiate validaiton method.An appropiate validaiton method.
* Results of the validation method or a future plan to obtain those results.

## Visualizations

Identify the location of at least three unique visulizations. They can additionally included here.

## User Guide

Include an enumerated (steps 1, 2, 3, etc.) guide to execute and use your application.

* Include instructions for downloaing and installing any necessary software or libraries.
* Provide an example of how the client should use the application.

# Reference Page

Following APA guidelines, include references for any cited works, e.g., (Author, year). References are not requried, and this page can be removed if no references are used. To cite sources used for code, you should include the referfences as code comments within the source code.