IMPaCT Intelligence

Workplace Violence (WPV) Prediction
Project Initiation Document

Prepared by: Carter Profico
Version: 1.0
Date: 02/10/2024
Sponsors: James Corbett (IMPaCT Intelligence), Arend van der Veen (WBAT Safety)

Document History

Version	Date	Updated By	Comments
1.0	02/10/2024	Carter Profico	
2.0	03/03/2024	Carter Profico	Revised all sections

Defined Terms

Term	Description
The Team	Refers to the Product Manager, Scrum Master, and Development Team (see Section 7, Organization and Governance)
The Sponsors	Refers to the individuals who serve as representatives from IMPaCT Intelligence

Table of Contents

1 Purpose

Workplace violence (WPV) is a prevalent issue in many industries, but it disproportionately affects those in healthcare settings. The goal of this project is to create a machine learning algorithm that is able to determine the risk of WPV incidents occurring based on the internal conditions of the hospital before they happen. This will allow security and other personnel to intervene in situations where WPV is likely to defuse the situation and prevent these incidents from happening in the first place.

2 Project Background

The inspiration for this project stems from the increasingly disproportionate acts of WPV in healthcare settings. The amount of WPV incidents are only increasing as time goes on, so having a method to detect these incidents before they happen will allow healthcare workers to remain safe in their workplace. According to the International Association for Healthcare Security and Safety's 2019 Healthcare Crime Survey, physical assault against healthcare workers in hospitals increased from 7.8 incidents per 100 beds in 2014 to 11.7 incidents per 100 beds in 2018.[1] Another source, the New Jersey Hospital Association says that New Jersey hospitals have reported a 14.6 percent increase in violent incidents between 2019 and 2021.[2] This is ignoring the fact that many WPV incidents go completely unreported, so the number of WPV incidents is almost certainly higher than we can see. The increasing number of WPV incidents calls for a better way to keep healthcare workers safe.

3 Scope

The primary goal of the Workplace Violence Prediction Model is to be able to develop a Machine Learning model to predict workplace violence in a hospital setting. The following lists will define the primary components to be completed, as well as the components we will consider out of scope.

3.1 In Scope

1	Creation, training and validation of a machine learning model
2	Utilize Hospital APIs to gather data
3	Process and clean gathered data
4	Generate predictions on gathered data
5	Create API routes that allows outside systems to use the WPV prediction model
6	Create a visual system that shows the current risk of WPV
7	Automatically send email notifications when there is high risk of WPV
8	Add user email addresses into the notification system
9	Allow for removal of users from the notification system

3.2 Out of Scope

1	Create individual models for each department
2	Add system that allows users to input when a WPV incident happens

4. Use Case Diagram

Below you will find a use case diagram for the project. This diagram provides a summary of the features, functions, and uses for the WPV prediction model. It also showcases the distinct actors and their individual actions that they are expected to take. An interesting addition here is the model itself as an actor in the diagram. This was included because of the large role the model itself will take, as it is an automatic monitoring system that works with little user input.
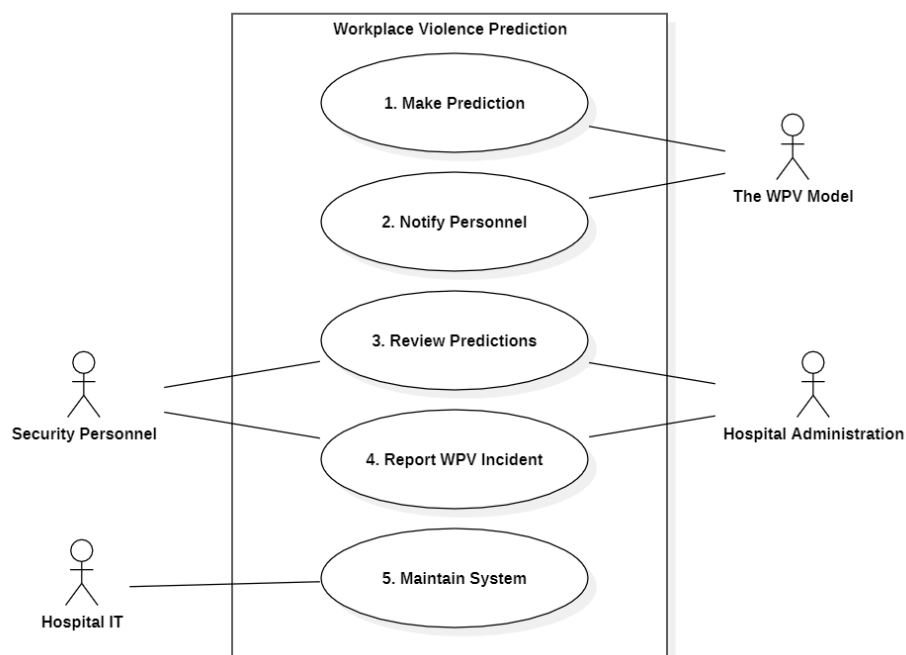
5   Assumptions and Constraints

5.1   Assumptions

There are a few assumptions the team will be working under. The first assumption is that the team will need to generate data that can then be used to train the model. There is a chance that real hospital data will be acquired, but it is

much safer to work under the assumption that there will not be real data. The

second assumption

5.2   Constraints

The largest constraint for this project is the time-window. There will be 5 sprints,

with each sprint being 2 weeks long. Those 10 weeks are the hard limit on the

time the team has to complete this project. Another constraint is the development

environment. The team will be using Python, Django, and MySQL as the primary

components of this project, as that is the standard development environment for

IMPaCT Intelligence.

6   Initial Project Plan

Sprint 0: Product Backlog

- The goals of this sprint are to organize the product backlog, write user stories and

  prioritize tasks. This sprint will prepare the development team for the work they will

  be doing throughout every following sprint. This sprint will also give the development

  team a headstart on understanding the frameworks they will be utilizing.

Sprint 1: Data Generation and Validation

- The goal of this sprint is to produce useful data. We do not have a source of data, so

  we must generate our own. After generation, we must also validate our data.

  Unvalidated data could cause our model to be inaccurate.

- Primary research topics will focus on the main predictors of WPV in hospitals and

  how to generate the necessary data to train and validate our model. They will also

include research into which library would be the best for the development team to utilize when creating the deep learning model.

Sprint 2: Deep Learning Model and Data Collection

- The goal of this sprint is to create a functioning deep learning model. Based on the research done in the previous sprint, the development team will utilize a Python library to create, train, and validate a deep learning model. They will also work on creating the system that will be used to integrate with hospital systems and gather the necessary data.
- Research must be done to understand how to connect with the various systems inside the hospital that will be used to collect data.

Sprint 3: Deep Learning Model and Data Collection

- The goal of this sprint is to continue working on the deep learning model and data collection. The model should make its predictions above 85% accuracy. The team will also continue working on the data collection systems, finishing up any necessary connections.
- Research must be done to understand how to connect with the various systems inside the hospital that will be used to collect data.

Sprint 4: API Creation and Integrating with Other Systems

- The development team will create an API that will allow other systems to integrate with the WPV Prediction Model. The API will allow other systems to push data to the model, which will then return a prediction. The development team will also focus this sprint on creating a method that will integrate the model into the Virtua hospital environment.
- The development team will research methods of making their API secure.

Sprint 5: User Interface and Email Notification System

- To make the WPV Prediction Model more intuitive, the development team will be creating a User Interface that will allow users to see the risk levels for each department. This UI will also include a way to add emails to the automated notification system. There will also be a place to log any WPV incidents. This will allow the model to grow over time as it gains a better dataset.

- The development team will need to research the libraries that can be utilized to create this UI system as well as the API necessary to automatically send out emails.

7 Organization and Governance

- The project's contributors are Aiden Touhill, Carter Profico, Avery Bobbitt, Joseph DiPietro, Christopher Duym and Anthony Ung.

- The project's sponsors are James Corbett (IMPaCT Intelligence) and Arend van der Veen (WBAT Safety).

- The Team consists of three main groups: The Product Owner, the Scrum Master, and the Development Team

  - The Product Owner, Aiden Touhill, is the primary communicator between the Team and the Sponsors. Aiden Touhill is also responsible for overseeing and delegating the project's tasks to the Development Team.

  - The Scrum Master, Carter Profico, is tasked with assisting the Development Team. Carter Profico will also ensure the Agile Scrum methodology is being followed.

  - The Development Team, consisting of Avery Bobbitt, Christopher Duym, Joseph DiPietro, and Anthony Ung, are the primary architects, designers, and coders for

the project. The Development Team follows the guidance set forth by the Sponsors and Product Owner and implements the project.

8 Communication Plan

- The team meets for fifteen-minute daily scrums three times a week at the beginning of the Software Engineering class.
- Sundays between 4:00 PM and 7:00 PM, Tuesdays between 5:00 PM and 6:15 PM, and Fridays between 11:00 AM and 1:30 PM are reserved for in-person meetings with the scrum team.
- Sprint reviews will be held every two Tuesdays at 2:00 for thirty minutes starting February 13th, not counting the week of March 11th.
- Any additional communication between the scrum team and the sponsors will be conducted by the product owner, Aiden Touhill, over email.

9 Risk Assessment

There are two primary risks for this project. The first is not having access to the data that is needed to properly train and validate the model. The core of any machine learning model is the data itself, so good data is a necessity. Without good data, the model will be inaccurate. This is a high likelihood and low severity risk. The high likelihood is because

	SEVERITY →		
LIKELIHOOD ↓	Low	Moderate	High
Low	Low	Low	Moderate
Moderate	Low	Moderate	High

gathering real-time data about hospital practices and employees is very difficult. To

mitigate this risk, the team plans on working under the assumption that data will not be

available and must be generated first. There are reputable sources with hospital data

available for free, which the team will utilize to generate the most accurate data possible.

The second primary risk is not having access to live hospital systems.. Without access to these systems, the development team would be unable to fully integrate the data collection and prediction system in a live environment, and must simulate any usage.. This is a high likelihood and low severity risk. The high likelihood is because it is difficult to gain access to a hospital's systems due to all of the protected information kept within. To mitigate this risk, the team will focus on creating an API that will be very flexible, to allow other systems to integrate with the WPV Prediction model

	SEVERITY →		
LIKELIHOOD ↓	Low	Moderate	High
Low	Low	Low	Moderate
Moderate	Low	Moderate	High
High	Moderate	High	High

10 Estimated Cost and ROI

10.1 Estimated Cost

Assuming our sponsors can attend every sprint review for 1.25 hours, they will

spend a total of 6.25 hours on the project in person or via teleconference.

Between the two sponsors, that equates to around $1500 worth of time. There

are no costs associated with the licensing of third-party software or equipment

needed, as everything the team is using is open-source. There are also no costs associated with the employment of the scrum team.

10.2 ROI

The ROI on this project is very hard to measure. There will be no monetary return on this project. While that may be true, the value of this product should not be understated. With a successful and accurate model, WPV incidents that are left unaddressed should drop dramatically. This will ensure a safer work environment for all hospital employees and patients. Not only will this create a safer work environment, it could even save lives. Hospitals are high intensity workplaces, where sometimes even the smallest mistake could mean the loss of a life. Time not spent dealing with WPV incidents is time that can be spent helping other patients.

11 References

1. Vellani KH. *The 2019 IAHSSF Healthcare Crime Survey*. IAHSS Foundation; 2019. Accessed March 3, 2024. https://iahssf.org/assets/2019-Healthcare-Crime-Survey-IAHSS-Foundation.pdf
2. *Hospital Workers Face a Growing Tide of Workplace Violence*. (n.d.). https://www.njha.com/media/699115/workplace-violence-bulletin-6-23-22-final.pdf