Inspiring Light Electronics Software Development Limited

I-Cremation

1 TABLE OF CONTENTS

2 PROJECT OVERVIEW

The goal of this project is aimed to automate the Wo Hop Shek Crematorium process to reduce the workload of cremation operations and improve the efficiency of cremation process by reducing the instability of performance of operators. To achieve the goal mentioned, we suggested to design a system utilizing A.I. technologies to monitor cremation scenes and give out optimized suggestions for the air inlets configuration. Operators can determine whether to adopt it or not. In the long run, the system is expected to fully take over the control of air inlets.

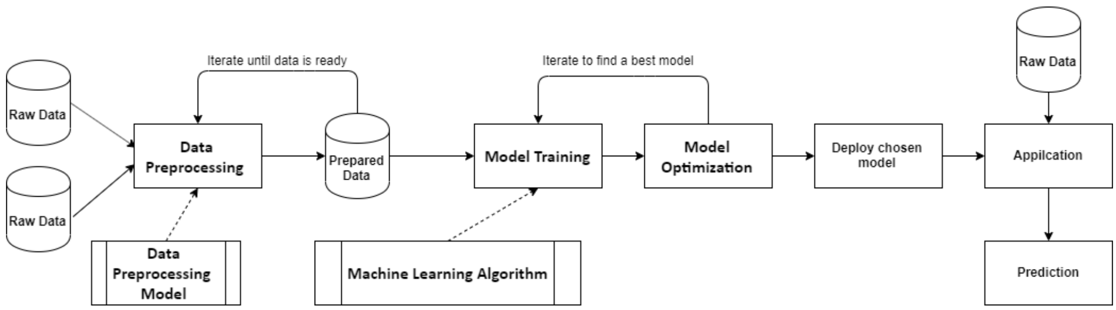
3 RECOMMENDED SOLUTION

3.1 SYSTEM ARCHITECTURE

The system would consist of two parts. The first part would be the A.I. training program and the second part would be the control program.

3.2 A.I. TRAINING PROGRAM

The objective of this program is to train an AI to simulate the operator decision making process and give out suggestions to the control program. The sections below are going to explain the principal of the AI Program and how it is going to be implemented:



3.2.1 Problem Identification

As the adjustment of air inlets from operations are continuously changing overtime, this machine leaning problem is induced as a regression problem. Given a specify feature $(\hat{x})$ to predict the configured value of 6 set of air inlet $(\hat{y})$ which $\hat{y} = [y_1 \quad \cdots \quad y_6]$, $y_i \in [0,100]$. In other words, the target is to figure out the hypothesis function $h_{\hat{w}}(\hat{x}) = \hat{y}$, where $\hat{w}$ represents the weight of the hypothesis function.

3.2.2 Data Acquisition & Exploratory Data Analysis

To train the A.I., following training data must be provided by EMSD, which include but not limited to:

1). Pre information data $(\widehat{x_p})$ (I.E. Age, Gender, Coffin materials, …)

2). Running information data ($\widehat{x_r}$) (I.E. Temperature, Air pressure, Video, Time, Pressure, Air inlets setting value before changed, …)

3). Operator decision data ($\hat{y}$) (I.E. Air inlets setting value after changed).

Since some of the environment features will not be considered by operators when making adjustment to the air inlet, those will be discarded after analysis.

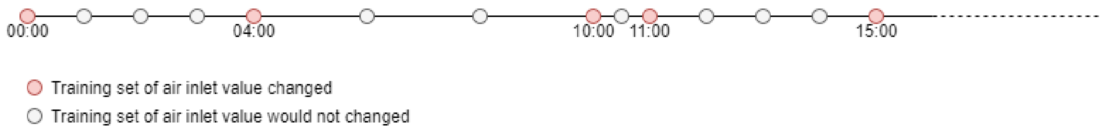
3.2.3 Data Preprocessing & Feature Engineering

Before feeding data to the A.I. program, several types of data have to be pre-processed due to the size and process power limitation.

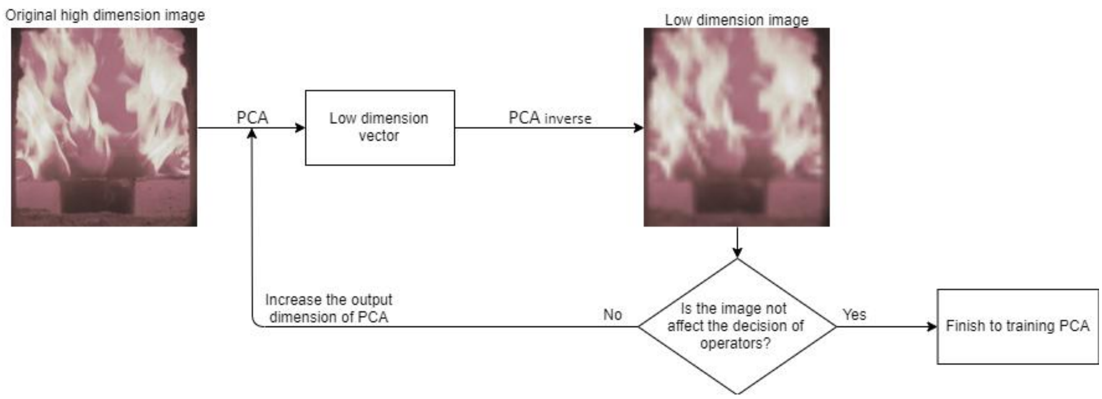
First, in order to measure the decision changes across a time series, a training set would be created whenever the operator adjusted the value of air inlets. As for reference, at least 3 more training set would be created between changes.

Second, to normalize the range of features, each feature would be scaled into $[0,1]$. So, it can increase the accuracy of the training model and decrease the training time.

An example of getting running information data



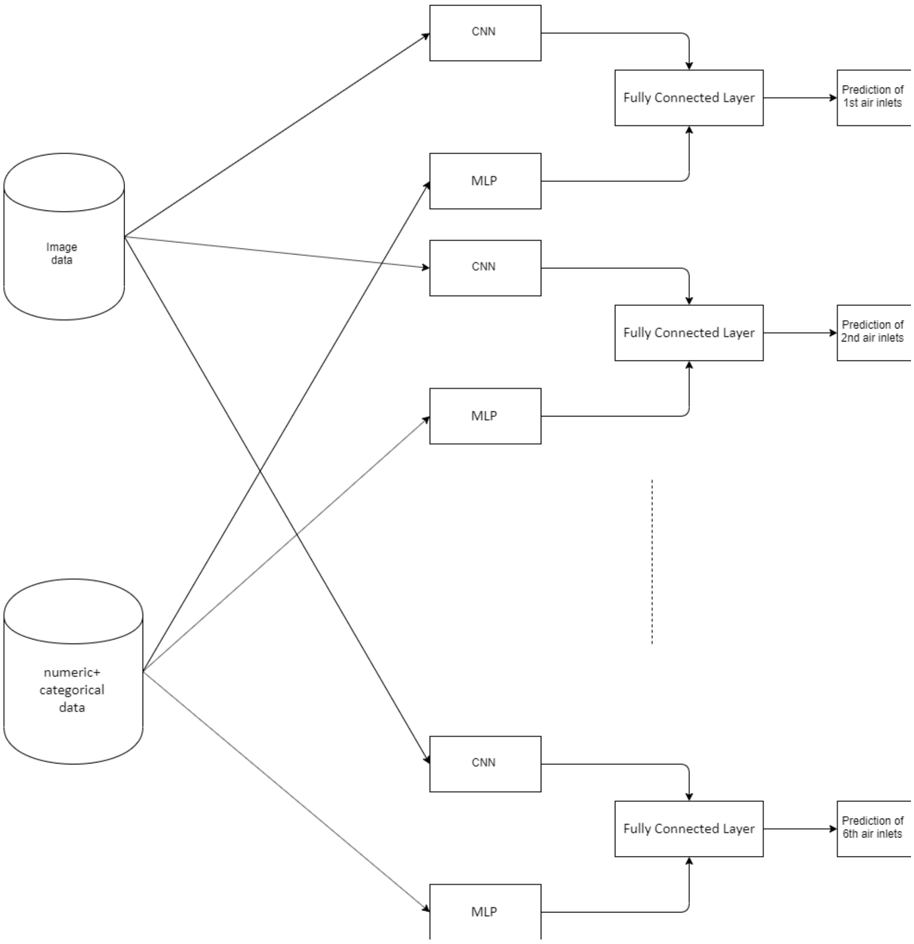
Since an image usually contains a lot of data (8million+ pixel), it would cause a underfit model and cost lots of training time. To deal with it, Principal Component Analysis (PCA) is a suitable method to reduce the dimension of the image data. This method would iterate it by different dimension to find a minimum dimension which would not affect the operator decision.



After going through all the procedures, many sets of training data $(\hat{x}, \hat{y})$, where $\hat{x} = (\widehat{x_p}, \widehat{x_r})$ would be obtained.

Finally, those training sets would be randomly divided into three categories where 60% is used for training, 20% for cross validation and 20% for testing.

3.2.4 Model Training



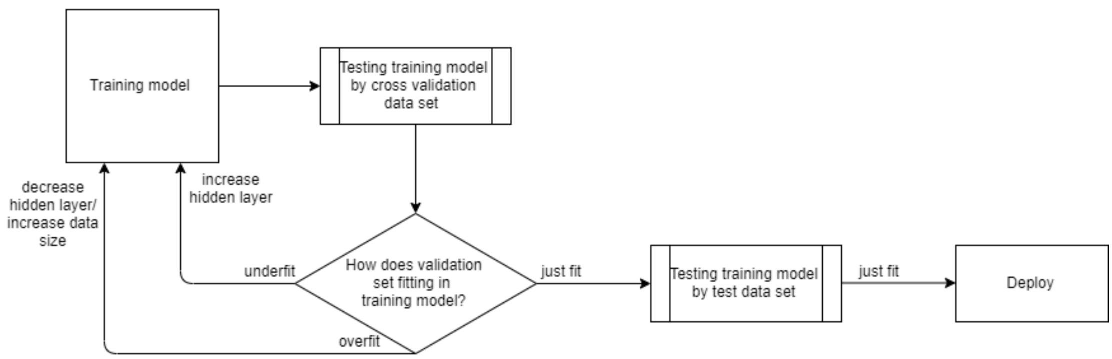
Since the property of the image data and numeric/categorical data is different, it is suggested to be separating these data for training. Convolutional Neural Network (CNN) is most commonly applied to analyze visual imagery, and Multilayer Perceptron (MLP) is a class of Artificial Neural Network (ANN) which can be applied to regression. The output of CNN and MLP would be passed into fully connected layer, which return the result of the prediction of a setting value of an specify air flap.

As the result of the model should be 6 values, therefore the model would be implemented for 6 times and trained separately. Each result of the training model represents each setting value of air flap respectively.

Mean Squared Error (MSE) would be used to be the cost function of the training model. In order to minimize the cost function, Gradient Descent would be used to find the local minimum of the cost function. The weight $(\hat{w})$ would be optimized, therefore the hypothesis function $h_{\hat{w}}(\hat{x}) = \hat{y}$ would be found.

3.2.5 Model Optimization

In order to find the suitable number of hidden layers, we would train the model iteratively by different number of hidden layers. After testing the model by cross validation set data, a "just fit" model would be found and avoided underfitting and overfitting.



3.2.6 Deployment & Future Evolvement

After it passed the test, if the error is controlled under a certain percentage, then it can be deployed and start emitting configuration values to the control program.

Furthermore, the A.I. can keep evolving and become more accurate and reliable. An online learning algorithm will be implemented and keep training the A.I. after deployment.

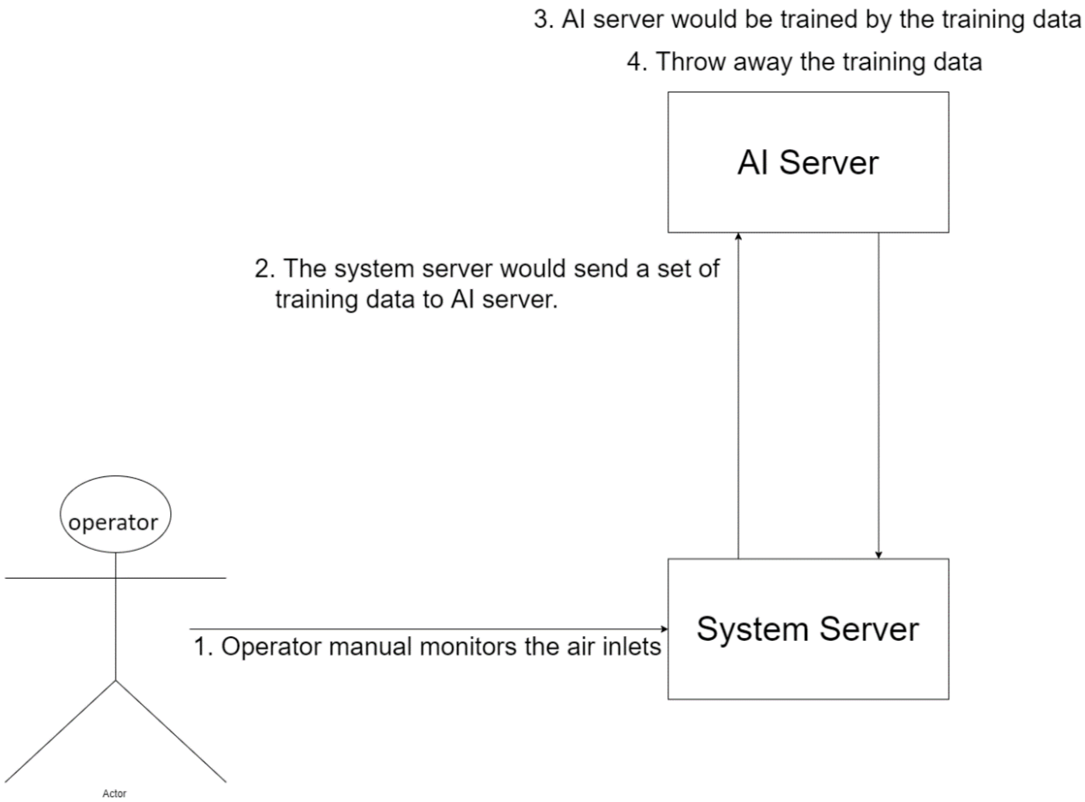
The process of training would be:

1. If no operator manually monitors the air inlets, which means the AI server can predict the value of the air inlet correctly, then we would not modify the hypothesis function.

2. When operator manual monitors the air inlets, the system server would send a set of training data $(\hat{x}, \hat{y})$ to A.I. server.

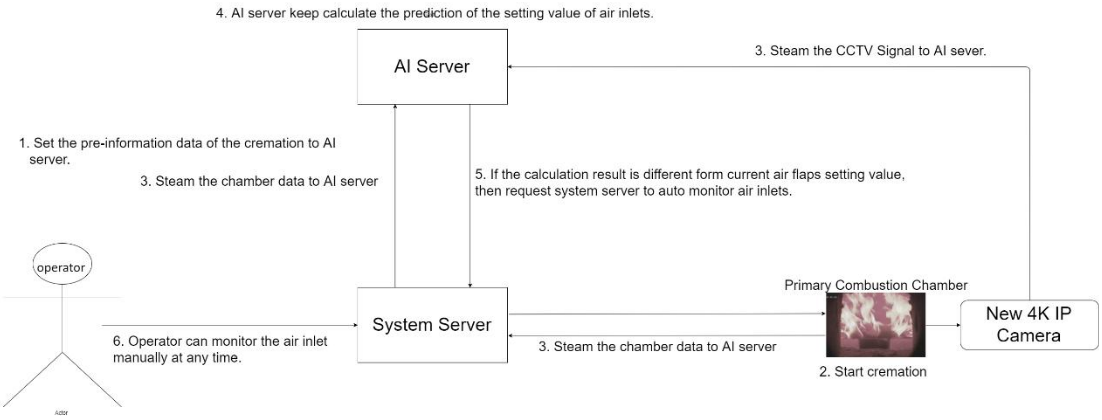
3. The AI server would be trained by the training data by Stochastic Gradient Descent (SGD), and modify the hypothesis function to fit this data.

4. After the A.I. system is trained, the training data $(\hat{x}, \hat{y})$ can be discarded.



3.3 CONTROL PROGRAM

This program would be served as a control panel for consuming the value emitted from the AI program and display the suggestions in a graphical user interface. It would be be able to take over the control of the air inlets of current cremation system if needed.



Estimated schedule

Task	Time (week)
Get Data & Exploratory Data Analysis	0.5
Data Preprocessing & Feature Engineering	1.5
AI Model development	3
Model Training & Model Optimization	3
AI server development	2
Front-end program development	1
Online learning program development	3
UAT and Trial Test	2
Total	16