Script:

We will first start off with a short technical explanation of what our team has come up during this hackathon before we embark on our demonstration. Given the freedom of such a wide problem statement, we have explored the different means Machine learning and Artificial Intelligence can benefit the society or the industry and we eventually settled upon the food industry. This is a common issue that is all around us and we often turn an blind eye to; the wastage of fresh produce. Oftentimes, produce that are overripe are not as visually appealing to consumers, causing them to be left on shelfs to spoil and they will be eventually thrown away. We propose having an additional screening process implemented by machine learning to identify over-ripen apples that allows for them to be set aside to be sold at a lower price to consumers with a lower tolerance for fruit preference or to consumers who are using these apples as an intermediate products such as baking which does not place a high priority on the apple’s appearance. This not only reduces the wastes generated at down-streamed supermarkets but enables these less-appealing apples to be sold, bringing about more revenue to the company.

Subsequently, we will be going through the different tools used as well as the rationale behind the choice of tools:

The first tool that we have utilized is machine learning and more specifically Convolutional Neural Network (CNN) models to classify fruits. For our prototype, the model is trained on classifying Ripe Apples, Rotten Apples and Over-Ripen Apples due to time and resources constraints. An issue faced in this aspect would be the collection of over-ripe apple datasets as most data sets available online are from both extremes; perfect apples or apples literally rotten down to the core. As a result of this issue coupled with the time constraints, we have decided to apply data augmentation strategies which we have drawn reference from research papers to increase the amount of training dataset the CNN model trains on. This allows the model to be well-developed as well as prevents the model from being overfitted onto the same dataset. Through image altering functions that rotate the images at various angles, adjust the brightness of images as well as blur the images, we are able to generate more data sets which are not identical. Additionally, this method of augmenting data helps us in our case whereby the images captured by production factories vary from location in terms of brightness, positioning and quality of camera used, as our CNN model is prepared for these data.

The second tool we used would be Edge Impulse. With production plants expected to process large amounts of produce at any time, it is crucial that this implemented strategy to reduce wastage does not influence their production rate, lowering productivity. The ability to easily train edge devices such as phones to with the actual trained machine learning model allows for a much faster response time as the computations will be evaluated on the devices themselves rather than being sent to a backend server, which may result in delays in responding and it may flood the network. Equipped with the capabilities to reduce the overall size and hence computational requirements of models uploaded into the hardware through various methods such as quantization, it produces similar accuracy rates yet at lower computational requirements as compared to centralized machine learning models.

The last tool used in our prototype would be the hosting of Edge Impulse over a website application. Having considered the needs of a production plant which may be handling different types of fruits, it may be more accurate for dedicated trained CNN models to be uploaded to the monitoring system rather than a general CNN model and having a server allows for easy deployment of these new updates to the various monitoring modules. Furthermore, it enables managers to have a better overview of the monitoring system through a consolidated platform which consolidates the number of fruits categorized under each label, allowing for better decisions to be made by the top management.

We will now proceed onto the demonstration of our prototype. This is a limited demonstration due to this being in the early developmental stage. Hope you will enjoy it!