Model

At first glance, the problem seems unsolvable using machine learning algorithm because we are not given the angle nor the velocity the projectiles were launched at. However, using my knowledge in physics, the velocity is calculated by distance over time when the time difference is small enough such as 10 ms. As a result, we can use the first two data point to calculate the angle and the velocity the projectile is launched at.

So I constructed such a model, $Y = W^T X$, where X contains the velocity, angle and time information of all 100 set of trajectories points, and Y contains all the location information (x_i, y_i). Such an equation can be solved using multi-dimension linear regression.

To be specific, the module I used is:

$$\begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} & x_n \\ y_0 & y_1 & x_2 & \cdots & y_{n-1} & y_n \end{bmatrix} = W^T \begin{bmatrix} v_x t_1 & v_x t_2 & \dots & v_x t_n \\ v_y t_1 & v_y t_2 & \dots & v_y t_n \\ t_1 & t_2 & \dots & t_n \\ t_1^2 & t_2^2 & \dots & t_n^2 \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

where v_x is the velocity along x axis, v_y is the velocity along y axis, and t_i is the time span the projectile spent and the current position. For the purpose of simplicity, I changed all time unit to be ms. Therefore, the least square solution W will be the coefficients we are going to use.

Why chose this model

Since all the data provided are trajectories, and from physics:

$$x = \vec{v}\, cos\theta\; t = \overrightarrow{v_x} t$$

$$y = \left((\vec{v}\, sin\theta) - a\right)t - \frac{1}{2} g t^2 + c = \left(\overrightarrow{v_y} - a\right)t - \frac{1}{2} g t^2 + c \; = \overrightarrow{v_y}\, t - at - \frac{1}{2} a\, t^2 + c$$

So, all the known parameters that concern the positions are $\overrightarrow{v_x}, \overrightarrow{v_y}, t, t^2, 1$

If the physics module is different, different parameters need to be chosen.

How to evaluate the model

Evaluate the model: construct the X matrix using velocity, degree and time information, then the data points of trajectory would be in the matrix $W^T X$

Testing the model: Since we are give 100 data sets, we can do cross-validation: use random 50 of them for training purpose and the other 50 of them for testing. Then repeat the process multiple times to find the best coefficients.

Predict projectiles launched at arbitrary angles and velocities

Yes, the model can predict projectiles at arbitrary angles and velocities well.

Assumptions made

I assumed that position is a linear function of $\overrightarrow{v_x}, \overrightarrow{v_y}, t, t^2, 1$

Will your approach/model change if we hadn't told you that the data was from a projectile?

Yes, it would change. If I wasn't provided the detail that the data points were from a projectile, I wouldn't have applied a physics module to the data points, and calculate the velocity of the projectile.

Did you refer to any relevant literature while solving this problem?

No, I didn't refer to any relevant literature while solving this problem.
My approach was, first I attempted to apply several machine learning algorithms I knew directly, but none of them lead to feasible results. I can't use supervised learning methods because the question is not a yes/no question, nor it is a classification problem. I can't use dimension reduction or clustering from unsupervised learning, because the dimension of each dataset are different. Also reinforcement learning doesn't work, because there's no action to make.
So I need to summarize a couple of features from those trajectories, and try to recover the trajectory points from those features.

If you were given enough time, how would you improve the model?

If given enough time, I would make it more generic. I would use more parameters to make the model higher dimension.