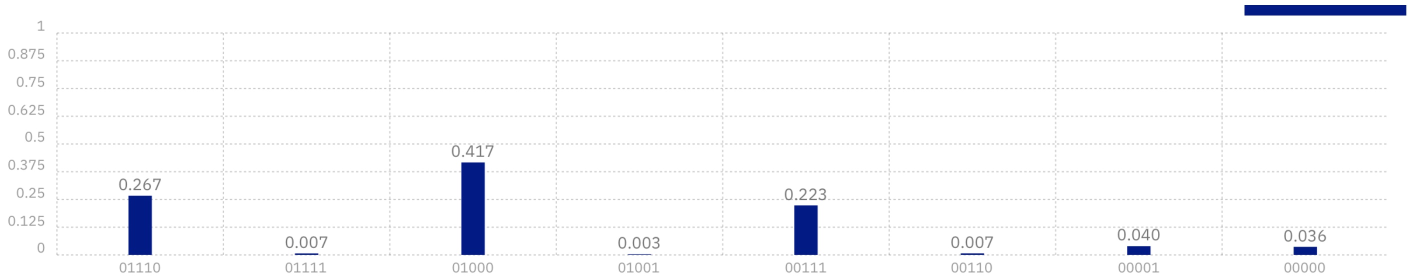

Figure 1: Classification of the input vector x'

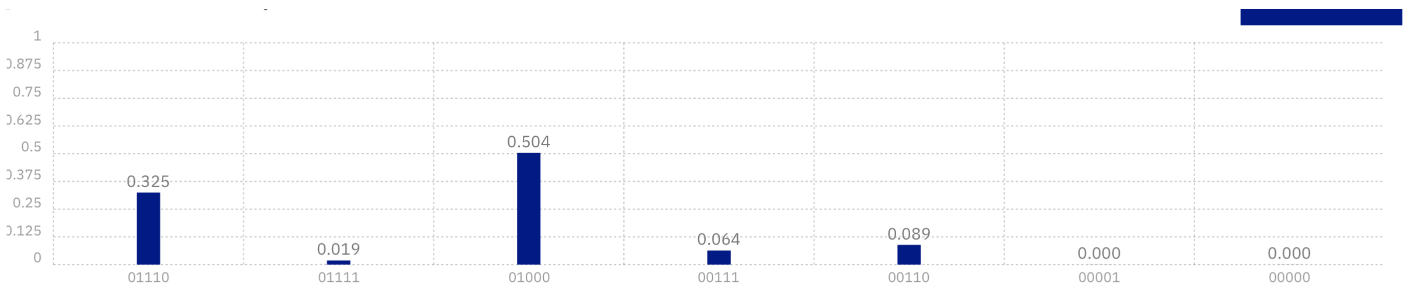

Figure 2: Classification of the input vector x''

Source code (specific to `ibmqx2`) from Mark Fingerhuth at
`https://github.com/markf94/ibmq_code_epl_119_60002`
copied here on Overleaf.

Note that IBM machine displays qubits in reverse order:

$$|q_4, q_3, q_2 = \text{class}, q_1, q_0 = \text{ancilla}\rangle$$

The algorithm includes a post-selection conditional on the ancilla qubit being 0. This means we need to discard results where the ancilla is 1 reducing the relevant space to quantum states $|01110\rangle, |01000\rangle, |00110\rangle, |00000\rangle$. This creates a normalizing constant:

$$Z = \Pr(|01110\rangle) + \Pr(|01000\rangle) + \Pr(|00110\rangle) + \Pr(|00000\rangle)$$

Classification of the input vector x'.

$$\begin{aligned}
Z &= 0.267 + 0.417 + 0.007 + 0.036 = 0.727 \\
\Pr(|c\rangle = |0\rangle) &= \frac{\Pr(|01000\rangle) + \Pr(|00000\rangle)}{Z} = \frac{0.417 + 0.036}{Z} = 0.623 \\
\Pr(\langle c| = \langle 1|) &= \frac{\Pr(|01110\rangle) + \Pr(|00110\rangle)}{Z} = \frac{0.267 + 0.007}{Z} = 0.377
\end{aligned}$$

Classification of the input vector x''.

$$\begin{aligned}
Z &= 0.325 + 0.504 + 0.089 + 0 = 0.918 \\
\Pr(|c\rangle = |0\rangle) &= \frac{\Pr(|01000\rangle) + \Pr(|00000\rangle)}{Z} = \frac{0.504 + 0}{Z} = 0.549 \\
\Pr(|c\rangle = |1\rangle) &= \frac{\Pr(|01110\rangle) + \Pr(|00110\rangle)}{Z} = \frac{0.325 + 0.089}{Z} = 0.451
\end{aligned}$$

1 Building the quantum circuit

The objective of the state preparation is to build the following quantum state:

$$|\mathcal{D}\rangle = \frac{1}{\sqrt{2MC}} \sum_{m=0}^{M-1} |m\rangle_\beta \big(|0\rangle_\alpha |\psi_{\mathbf{x}'}\rangle_\gamma + |1\rangle_\alpha |\psi_{\mathbf{x}^m}\rangle_\gamma\big) |y^m\rangle_\delta \tag{1}$$

containting $M = 2$ training examples and where $\alpha, \beta, \gamma, \delta$ respectively stand for the ancilla, index, data and class qubits. For clarity and hints about how to actually prepare the state it is useful to expand as:

$$|\mathcal{D}\rangle = \frac{1}{\sqrt{2MC}}\left[|0\rangle_\beta\big(|0\rangle_\alpha|\psi_{\mathbf{x}'}\rangle_\gamma + |1\rangle_\alpha|\psi_{\mathbf{x}^0}\rangle_\gamma\big)|y^0\rangle_\delta + |1\rangle_\beta\big(|0\rangle_\alpha|\psi_{\mathbf{x}'}\rangle_\gamma + |1\rangle_\alpha|\psi_{\mathbf{x}^1}\rangle_\gamma\big)|y^1\rangle_\delta\right]$$

Vectors to be loaded are:

$$\begin{aligned}
|\psi_{\mathbf{x}'}\rangle &= -0.549|0\rangle + 0.836|1\rangle \\
|\psi_{\mathbf{x}^0}\rangle &= |1\rangle \\
|\psi_{\mathbf{x}^1}\rangle &= 0.789|0\rangle + 0.615|1\rangle
\end{aligned}$$

The first step consists in bringing the ancilla and the index qubits in a superposition state. This allows to load the first input as entangled with both ancilla and index effectively creating 2 copies of the input vector:

$$\begin{aligned}
|\mathcal{D}\rangle &= |0\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha \\
&\sim |0\rangle_\gamma\big(|0\rangle + |1\rangle\big)_\beta\big(|0\rangle + |1\rangle\big)_\alpha
\end{aligned}$$

The next step is to load the test vector $\mathbf{x}'$. In this case, we want to entangle the vector only in the case that the ancilla qubit is $|1\rangle$.

$$\begin{aligned}
\mathrm{CR}_y(\alpha;\theta,\gamma) &\sim |0\rangle_\gamma\big(|0\rangle + |1\rangle\big)_\beta|0\rangle_\alpha + \left(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\right)_\gamma (|0\rangle + |1\rangle)_\beta\,|1\rangle_\alpha \\
&\sim |0\rangle_\gamma\big(|0\rangle + |1\rangle\big)_\beta|0\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma\,(|0\rangle + |1\rangle)_\beta\,|1\rangle_\alpha \\
&\sim (|0\rangle + |1\rangle)_\beta\left(|0\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|1\rangle_\alpha\right)
\end{aligned}$$

where:

$$\theta \approx 2\,\mathrm{arctan}_2(0.835754, -0.549104) \approx 4.304$$

$$\begin{aligned}
X_\alpha &\sim (|0\rangle + |1\rangle)_\beta\left(|0\rangle_\gamma|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha\right) \\
&\sim |0\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |0\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + (|0\rangle + |1\rangle)_\beta\,|\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha \\
\text{Toffoli}(\alpha,\beta;\gamma) &\sim |0\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |1\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + (|0\rangle + |1\rangle)_\beta\,|\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha \\
&\sim |0\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + (|0\rangle + |1\rangle)_\beta\,|\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha \\
X_\beta\text{ cute trick...} &\sim |0\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + (|0\rangle + |1\rangle)_\beta\,|\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha
\end{aligned}$$

The trick moves the first training vector $|\psi_{\mathbf{x}^0}\rangle$ to the $|0\rangle$ state of the index qubit β. This allows us to now apply another Toffoli gate to load the second vector into the $|1\rangle$ state of the index and doesn't change anything on the test vector $|\psi'_{\mathbf{x}}\rangle$ since it is already entangled with a index qubit β that is in a superposition state.

$$\begin{aligned}
X_\beta\text{ cute trick...} &\sim |0\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + (|0\rangle + |1\rangle)_\beta\,|\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\alpha \\
&\sim |0\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi'_{\mathbf{x}}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + |\psi'_{\mathbf{x}}\rangle_\gamma|1\rangle_\beta|0\rangle_\alpha \\
\text{Toffoli}(\alpha,\beta;\gamma) &\sim |1\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|1\rangle_\beta|0\rangle_\alpha \\
\mathrm{CR}_y(\beta;-\theta,\gamma) &\sim (R_y(-\theta)|1\rangle)_\gamma\,|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + (R_y(-\theta)|\psi_{\mathbf{x}'}\rangle)_\gamma\,|1\rangle_\beta|0\rangle_\alpha \\
\text{Toffoli}(\alpha,\beta;\gamma) &\sim \left(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\right)_\gamma |1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + (R_y(\theta)|\psi_{\mathbf{x}'}\rangle)_\gamma\,|1\rangle_\beta|0\rangle_\alpha
\end{aligned}$$

The trick above is that $(R_y(-\theta)|\psi_{\mathbf{x}'}\rangle)_\gamma\,|1\rangle_\beta|0\rangle_\alpha$ is unaffected by the second Toffoli gate since the ancilla is in state $|0\rangle$.

$$\begin{aligned}
\mathrm{CR}_y(\beta;\theta,\gamma) &\sim \left[R_y(\theta)\left(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\right)\right]_\gamma |1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + (R_y(\theta)R_y(-\theta)|\psi_{\mathbf{x}'}\rangle)_\gamma\,|1\rangle_\beta|0\rangle_\alpha \\
&\sim (\cos\theta|0\rangle + \sin\theta|1\rangle)_\gamma\,|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + (|\psi_{\mathbf{x}'}\rangle)_\gamma\,|1\rangle_\beta|0\rangle_\alpha \\
&\sim |\psi_{\mathbf{x}^1}\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}^0}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + |\psi_{\mathbf{x}'}\rangle_\gamma|1\rangle_\beta|0\rangle_\alpha
\end{aligned}$$

where:

$$\theta \approx \mathrm{arctan}_2(0.61489, 0.78861) \approx 0.662$$

We are now in a position to introduce the class qubit δ. We should simply condition it to be in state $|1\rangle$ when the index qubit β is also in state $|1\rangle$ and to leave it to $|0\rangle$ otherwise.

$$\begin{aligned}
& \sim & & \big(|\psi_{\mathbf{x^1}}\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x^0}}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha + |\psi_{\mathbf{x'}}\rangle_\gamma|1\rangle_\beta|0\rangle_\alpha\big)|0\rangle_\delta \\
\mathrm{CX}(\delta;\beta) & \sim & & \big(|\psi_{\mathbf{x^0}}\rangle_\gamma|0\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\beta|0\rangle_\alpha\big)|0\rangle_\delta + \big(|\psi_{\mathbf{x^1}}\rangle_\gamma|1\rangle_\beta|1\rangle_\alpha + |\psi_{\mathbf{x'}}\rangle_\gamma|1\rangle_\beta|0\rangle_\alpha\big)|1\rangle_\delta \\
& \sim & & |0\rangle_\beta\big(|\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x^0}}\rangle_\gamma|1\rangle_\alpha\big)|0\rangle_\delta + |1\rangle_\beta\big(|\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x^1}}\rangle_\gamma|1\rangle_\alpha\big)|1\rangle_\delta \\
& = & & \frac{1}{\sqrt{2M}}\Big(|0\rangle_\beta\big(|\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x^0}}\rangle_\gamma|1\rangle_\alpha\big)|0\rangle_\delta + |1\rangle_\beta\big(|\psi_{\mathbf{x'}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x^1}}\rangle_\gamma|1\rangle_\alpha\big)|1\rangle_\delta\Big)
\end{aligned}$$

Now that the state has been prepared we can finally run the algorithm which consists in applying a single Hadamard gate to the ancilla qubit.

$$\begin{aligned}
H(\alpha) & = \frac{1}{2\sqrt{M}}\Big(|0\rangle_\beta\big(|\psi_{\mathbf{x'}+\mathbf{x^0}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x'}-\mathbf{x^0}}\rangle_\gamma|1\rangle_\alpha\big)|0\rangle_\delta + |1\rangle_\beta\big(|\psi_{\mathbf{x'}+\mathbf{x^1}}\rangle_\gamma|0\rangle_\alpha + |\psi_{\mathbf{x'}-\mathbf{x^1}}\rangle_\gamma|1\rangle_\alpha\big)|1\rangle_\delta\Big) \\
& = \frac{1}{2\sqrt{M}}\sum_{m=0}^{M-1}|m\rangle_\beta\big(|0\rangle_\alpha|\psi_{\mathbf{x'}+\mathbf{x}^m}\rangle_\gamma + |1\rangle_\alpha|\psi_{\mathbf{x'}-\mathbf{x}^m}\rangle_\gamma\big)|y^m\rangle_\delta
\end{aligned}$$

where $|\psi_{\mathbf{x'}\pm\mathbf{x}^m}\rangle = |\psi_{\mathbf{x'}}\rangle \pm |\psi_{\mathbf{x}^m}\rangle$. Now we can perform a conditional measurement based on the value of the ancilla qubit:

$$\begin{aligned}
\Pr(|\alpha\rangle = |0\rangle_\alpha) & = \frac{1}{4M}\left(|\mathbf{x'}+\mathbf{x^0}|^2 + |\mathbf{x'}+\mathbf{x^1}|^2\right) \\
\Pr(|\alpha\rangle = |1\rangle_\alpha) & = \frac{1}{4M}\left(|\mathbf{x'}-\mathbf{x^0}|^2 + |\mathbf{x'}-\mathbf{x^1}|^2\right)
\end{aligned}$$

TODO: Maybe remove completely the notation $|\psi_{\mathbf{x}}\rangle_\gamma$ to simply $|\mathbf{x}\rangle_\gamma = \mathbf{x}_0|0\rangle_\gamma + \mathbf{x}_1|1\rangle_\gamma$. It would also help explain better the final superposition with Hadamard gate.

Therefore if the ancilla is measured $|\alpha\rangle = |0\rangle$, the post-measurement quantum state is now given by:

$$|\mathcal{D}\rangle_{|\alpha\rangle=|0\rangle_\alpha} = \frac{1}{2\sqrt{M\Pr(|\alpha\rangle = |0\rangle_\alpha)}}\sum_{m=0}^{N-1}|m\rangle_\beta|\mathbf{x'}+\mathbf{x}^m\rangle_\gamma|y^m\rangle_\delta \tag{2}$$

This normalized state can be viewed as a weighted sum:

$$\begin{aligned}
1 & = \frac{1}{4Mp_{\mathrm{acc}}}\sum_m|\mathbf{x'}+\mathbf{x}^m|^2 \\
1 & = \frac{1}{Mp_{\mathrm{acc}}}\sum_m\left(1-\frac{1}{4}|\mathbf{x'}-\mathbf{x}^m|^2\right) \\
1 & = \underbrace{\frac{1}{Mp_{\mathrm{acc}}}\sum_{m|y^m=0}\left(1-\frac{1}{4}|\mathbf{x'}-\mathbf{x}^m|^2\right)}_{\Pr(\tilde{y}=0)} + \underbrace{\frac{1}{Mp_{\mathrm{acc}}}\sum_{m|y^m=1}\left(1-\frac{1}{4}|\mathbf{x'}-\mathbf{x}^m|^2\right)}_{\Pr(\tilde{y}=1)}
\end{aligned}$$

where:

$$\Pr(\tilde{y}=0) = \frac{1}{4Mp_{\mathrm{acc}}}\sum_{m|y^m=0}\left(1-\frac{1}{4}|\mathbf{x'}+\mathbf{x}^m|^2\right) \tag{3}$$

Clearly, if the more the input vector is "closer" to the training input labeled as belonging to class 0, the more the value of $\Pr(\tilde{y}=0)$ rises. Because of the normalization of the post-measurement quantum state above, it is clear that if we find $\Pr(\tilde{y}=0) > 0.5$, the prediction is 0 otherwise we should predict class 1. This is achieving exactly the same effect as the original classifier. Running the quantum circuit gives us an experimental value for $\Pr(\tilde{y}=0)$ and the LHS gives us confirmation that this value indeed corresponds to the classification task we defined originally.