Exercise 5: Hello tf2

1. Watch the videos for tf.keras crash course

video part 1 and part 2
https://www.youtube.com/watch?v=UYRBHFAvLSs
https://www.youtube.com/watch?v=uhzGTijaw8A

and study **first_steps_with_tensorflow2.ipynb** (pay close attention to gradient tape)

2. Implement the Mish activation function

https://arxiv.org/abs/1908.08681
and use it instead of the ReLU activation function in **first_steps_with_tensorflow2.ipynb**

3. Proof the statement in the last section of the slides (see slides for hints)

- **Proof that:**

$$v_\pi(s) \geq v^*(s) - \frac{2\epsilon}{1-\gamma}$$

our policy performance

Given:

$$\max_{s,a} |q_\theta(s,a) - q^*(s,a)| \leq \epsilon$$

$$\pi(s) = \operatorname{argmax}_a q_\theta(s,a) \text{ (our approx.)}$$

$$\pi^*(s) = \operatorname{argmax}_a q^*(s,a)$$

$$q^*(s,a) = q_{\pi^*}(s,a)$$