1. Name each gate of an LSTM and give a short description of what it does:
   - **Gate 1**
     
     **Solution:** Forget gate: controls how much of the previous $c_{t-1}$ makes it into $c_t$. ($c_t = f_t \odot c_{t-1} + ...$). If the forget gate is 1, all of the previous cell state makes it into this cell state. If 0, none of it.
   - **Gate 2**
     
     **Solution:** Input gate: controls how much of the calculation based on $x_t$ and $h_{t-1}$ makes it into $c_t$. (0 means none, 1 means all.)
   - **Gate 3**
     
     **Solution:** Output gate: controls how much of $c_t$ makes it into $h_t$. (0 means none, 1 means all.)

2. Can the gates have different values at different timesteps $i$ and $j$ if the inputs at those timesteps, $x_i$ and $x_j$ are identical? Why or why not?

   **Solution:** A gate can have different values at different timesteps because it depends not only on $x_t$ but also on $h_{t-1}$ which may be different at timestep $i$ and timestep $j$. This question was not entirely clear: of course the different gates can have different values (even at the same timestep). For a given gate, though, can its value be different at different timesteps with the same input?