Question: True or False

(a) Contextual bandits - taking an action can result in change of state.

False There is only a single state, no state change is involved.

(b) LOOCV - a method for preventing overfitting.

True Cross validation uses a data point as training data & validation data (but not at the same time). Validation set error allows selecting a model that does not overfit the data.

(c) Q-Learning - data must come from the optimal policy.

False Q-Learning is an off-policy method.

(d) Learning - 80% class^n accuracy on training set & 20% " " " validation set] classifier is overfitting

True

(e) Value Iteration - Bellman update for one state $O(|S||A|)$

True

(f) TD-learning - needs a transition model of the MDP to be provided.

False TD learning is model free.

(g) Gradient descent with momentum - uses second derivative.

False

(h) Perceptron - no classification mistakes, algorithm has converged

True If output is correct then the weights are not changed.

(i) By using MCV & LCV heuristics, every CSP can be solved in time linear in the no of variables.

False

(j) UCS will never expand more nodes than A*-Search

False