Project 5 Questions

Instructions

- 3 questions.
- Write code where appropriate; feel free to include images or equations.
- **We do NOT expect you to fill up each page with your answer.** Some answers will only be a few sentences long, and that is okay.

Questions

Q1:

(a) Explain these common terms in machine learning in your words:

(i) Bias

(ii) Variance

(b) Define these terms in the context of evaluating a classifier:

(i) Overfitting

(ii) Underfitting

(c) Does bias and variance have any impact on overfitting and underfitting. Can you describe a brief real-world scenario where you can find/observe this?

Please answer overleaf.

A1: Your answer here. Uncomment the stencil below and fill in your solution.

(a) (i) Bias is the square of the average deviation of an estimator from the ground truth

(ii) Variance denotes is the variance of the estimates, i.e., the expected squared deviation from the estimated mean

(b) (i) Overfitting makes a classifier model complex. The model overfits the training data and gives high accuracy rate for training set but gives low accuracy rate for test set.

(ii) Underfitting makes the model simple. If we get low accuracy rate for both training and test set, then we can guess that the model is suffering from underfitting and we need more dataset to train the model.

(c) Overfitting increases bias and lowers variance. Underfitting lowers bias and increases variance.

Q2: Suppose you had to test the selective search algorithm on an image from pedestrian detection dataset (for example: an image taken from traffic camera as shown below), do you think that selective search algorithm will suggest bounding boxes over person(s).

(a) If yes, what is the justification in your words for this successful behavior of the algorithm? If not, then can you think in which cases does it fail? Can you suggest one way to improve or modify the approach?

A2: Your answer here.

(a) No, the image does not have enough colour variations. As a result, it does not give proper segmentation from Felzenszwalb algorithm. Maybe pre-processing the image by increasing the contrast would help.

Q3: If you were to apply selective search algorithm to detect interesting regions for skin diseases, how do you think the algorithm would have performed?

(a) Which among the four similarity (color, texture, size and shape) do you think contributes more to automatic detection of interesting regions in this case?

(b) Combining what you understood so far, what do you think "objectness" means and do you agree that selective search algorithm inherently finds regions of interest based on "objectness"?

Please answer overleaf.

A3: Your answer here.

(a) colour

(b) Objectness is essentially a measure of the probability that an object exists in a proposed region of interest. If we have high objectness, this means that the image window likely contains an object. This allows us to quickly prune out proposed image windows that do not contain any objects. Yes, selective search algorithm inherently finds regions of interest based on "objectness"

Feedback? (Optional)

Please help us make the course better. If you have any feedback for this assignment, we'd love to hear it!