1.

2. What is the difference between "model bias" and "model variance"?

(i). Why is a high bias, low variance classifier undesirable?

(ii). Why is a low bias, high variance classifier (usually) undesirable?

Bias: propensity of a classifier to systematically produce the same errors. $E[g(x) - f(x)]$ (average model approx error over all possible training sets)

If it doesn't produce error / produces different kinds of errors ⇒ unbiased (e.g. predict too many instances as the majority class)

Variance: propensity of a classifier to produce different classifications using different training set. (randomly sampled from same population)

Measure of the inconsistency of the classifier, from training set to training set.

$$E[\{f(x) - E(f(x))\}^2]$$

(i) High bias & low variance

⇒ Consistently wrong.

(ii) Low bias & high variance

low bias ⟶ may be correct predictions.

high variance ⟶ difficult to be certain about the performance of the classifier

If high variance, ER may be low on one set of data, and high on another set (not generalised)

2.

~~J~~. Describe how validation set, and cross-validation can help reduce overfitting?

models usually have hyperparameter(s) ⟶ control model complexity

find best values ⟶ to achieve best predictive performance on new data.

(may also consider a range of different types of models ⇒ find best one)

performance on training data : not a good indicator on unseen data.
(may be overfitting)

2 ways :

① Validation set: we train models on training set, compare them on independent data (val set) ⇒ select best one.
evaluate the final model with test set.

② CV: If data is limited & want good models
⇒ use as much of the available data as possible for training.
⇒ small validation set ⇒ Use CV

3.

~~A~~. Why ~~ensembling~~ bagging reduces model variance?

Ensembling :

Z_1, Z_2, Z_3 : models (assume equal var)

$$\mathrm{Var}\left(\frac{1}{N}\sum_i Z_i\right) = \frac{1}{N^2}\mathrm{Var}\left(\sum_i Z_i\right)$$
$$= \frac{1}{N^2}\left[\mathrm{Var}(Z_1) + \mathrm{Var}(Z_2) + \cdots\right]$$
$$= \frac{1}{N^2}\cdot N \cdot \mathrm{Var}(Z_1)$$
$$= \frac{\mathrm{Var}(Z_1)}{N} \quad \text{(smaller)}$$