

Fiche d'auto-évaluation 13

February 3, 2022

1. Qu'est-ce l'importance de permutation (Permutation Importance) et comment la calcule-t-on?
2. L'importance de permutation, donne-t-elle des informations sur le jeu de données (informations qui ne dépendent pas du modèle supervisé utilisé) ou sur le modèle supervisé?
3. Que prédit le Kaplan-Meier estimator? Est-il utile pour classer un type d'anomalie? Ou pour avoir une idée de la probabilité de faillite d'une machine?
4. Qu'est-ce que la fonction de survie (survival function) $S(t) = \Pr(T > t)$ signifie?
5. La fonction de survie $S(t) = \Pr(T > t)$ est-elle monotone? Si oui, décroissante ou croissante? Quelle est sa valeur pour $t = 0$? Et pour $t \rightarrow +\infty$?
6. Définissez la fonction de hasard (hazard function).
7. Donnez des exemples de modèles que nous avons vus pour estimer la fonction de survie et de hasard? Sont-ils des modèles supervisés ou non-supervisés?
8. Dans l'analyse de maintenance prédictive, donnez des exemples d'exemples censurés.
9. Qu'est-ce l'index de concordance (concordance index)? Créez un petit exemple où vous avez une suite de 3 événements et un modèle qui prédit l'ordre de ces 3 événements. Calculez l'index de concordance.
10. Vous êtes ingénieur de production dans une usine qui travaille à des processus à très hautes températures. Vous y avez installé deux modèles de capteurs et vous avez collecté les données relatives à leur fonctionnement. En particulier, vous avez noté quand les capteurs se cassent. Sur la base de ces informations historiques, vous voulez comparer leur qualité, en terme de résistance (probabilité de survie). Quelles courbes pouvez-vous dessiner à ce propos, parmi celles vues en cours? Dessinez deux courbes hypothétiques, en imaginant qu'un modèle de capteur est plus résistant que l'autre.

11. Quelle est la différence entre mesures de réseaux passives et actives? Quels sont les avantages et les inconvénients?