Machine learning

Gestion de l'incertain

Travaux Dirigés

1 Filtrage bayésien : système de surveillance du trafic aérien

On considère dans un premier temps un système de surveillance du trafic aérien reposant entièrement sur le système GPS : chaque avion réalise l'acquisition de sa position GPS avant de la transmettre par voie hertzienne au centre de contrôle.

▷ **Q1** : Proposez un système pour suivre en temps-réel la position d'un avion tout en intégrant l'incertitude sur les mesures. On précisera les équations mises en jeu.

Le centre de contrôle est équipé d'un radar de surveillance de l'espace aérien, monté sur une plateforme mécanique en rotation dont la fréquence de révolution est de 0,5 Hz. Le principe d'un radar consiste à émettre des ondes électromagnétiques dans une direction du ciel. Si les ondes rencontrent au cours de leur propagation un objet volant, une partie de l'énergie de l'onde est réfléchie en direction de l'antenne émétrice. La détection de cet écho au niveau du radar permet de mesurer approximativement la direction (en cap et élévation) et la distance de l'objet dans le référentiel du radar. On observe que l'incertitude sur l'angle de reception est fixe alors que celle sur la distance estimée est proportionnelle à la distance réelle de l'objet à l'antenne.

▷ **Q2** : Modifiez votre système de suivi en y intégrant les informations du radar.

2 Classification bayésienne : évasion fiscale

Le service des contrôles fiscaux veut améliorer son efficacité en focalisant ses contrôles sur les personnes les plus susceptibles de frauder. Le service dispose d'une base de données décrivant des millions de contribuables. Chaque personne y est décrite par des dizaines d'attributs : catégorie socio-professionnelle, type et montant des revenus déclarés, situation familiale, liste des biens connus : immobiliers (valeur, localisation, etc), voiture (type, valeur), etc. Une centaine de milliers de contribuables ont été controlés dans le passé et pour une dizaine de milliers d'entre eux, un redressemement fiscal a été réalisé et notifié dans la base de données.

▷ **Q3** : On suppose dans un premier temps que tous les attributs continus ont été discrétisés. On suppose également que les autorités veulent contrôler la totalité des personnes soupconnées de fraude fiscale. Aidez les autorités en concevant une méthode pour déterminer les contribuables à contrôler.

▷ **Q4** : Comment intégrer les attributs numériques tels que les revenus déclarés, la valeur du patrimoine, etc

▷ **Q5** : Quels sont les points forts et faibles de la méthode ? Le problème n'est-il pas biaisé ? Comment prendre en compte ce biais ?

▷ **Q6** : On suppose maintenant que les autorités peuvent contrôler jusqu'à cent personnes par jour. Comment cela affecte-il la méthode ?

▷ **Q7** : On suppose que le contrôle fiscal d'un contribuable coûte en moyenne 1000 € à l'état et que la découverte d'une fraude rapporte en moyenne 20000 €. Comment intégrer cette information dans votre modèle ?

▷ **Q8** : On considère les attributs suivants : le salaire (S), métier (M), secteur d'activité (A), patrimoine (P), le fait d'avoir déjà fait l'objet d'un redressement (R) et la variable cible (C) (fraudeur ou non). En outre on introduit le capital financier (F) qui n'est que partiellement observable. Proposez un modèle plus réaliste que Naive Bayes pour tenir compte des dépendances entre variables. Adaptez si nécessaire le modèle pour vérifier les relations d'indépendance suivantes :

1. $\forall X, \forall Y \neq R,\ X \not\perp Y$
2. C et P n'ont pas de fils.
3. $\forall X \neq C,\ R \perp X$
4. $\forall X \notin \{A, C, F, R\},\ X \perp C \mid F, A, R$
5. $\forall X \notin \{F, P\},\ X \perp P \mid F$
6. $\forall X \notin \{A, S, M\},\ X \perp M \mid A, S$

Quel est le principal inconvénient d'un tel modèle par rapport à Naive Bayes ?

▷ **Q9** : On suppose que M et A peuvent prendre chacune 30 valeurs, que S et P suivent en première approximation des distributions normales et que la valeur moyenne et l'écart-type de P sont des fonctions quadratiques de S. Calculez le nombre de paramètres nécessaires et comparez avec Naive Bayes. Concluez.

3 Bandit bayésien (3 point(s))

Dans cet exercice, on considère le cadre usuel des bandits stochastiques à K bras. Chaque bras $i \in \{1, \ldots, K\}$ est associé à une mesure de probabilité inconnue ν_i. Ici, on supposera de plus que cette mesure suit une *loi de Bernoulli*. A chaque instant $t = 1, 2, \ldots$, le joueur choisi un bras $I_t \in \{1, \ldots, K\}$ (en fonction des choix et observations passés) et reçoit une récompense $X_{I_t,t}$ tirée selon ν_{I_t}, indépendamment du passé. On note $\mu_i = \mathbb{E}[X_{i,t}]$ l'espérance du $i^{\text{ème}}$ bras et on définit $\mu_* = \max_{1 \leq i \leq K} \mu_i$ et $i_* \operatorname{argmax}_{1 \leq i \leq K} \mu_i$.

Ici, plutôt que de considérer une approche fréquentiste et dériver un algorithme à partir d'une inégalité de concentration (par exemple UCB à partir de l'inégalité de Hoeffding), on considère une approche bayésienne. L'idée est de paramétrer chaque bras par un paramètre caractérisant la loi de Bernoulli, ce qui donne la vraisemblance, d'imposer une distribution a priori sur ce paramètre et de calculer la distribution a posteriori en fonction des récompenses observées.

▷ **Q10** : Quelle est la distribution conjuguée d'une loi de Bernoulli ? Notons $x_{i,1}, \ldots, x_{i,s}$ les récompenses observées après avoir tiré s fois le bras i. Après avoir expliciter le paramètre associé (au bras, donc à la loi de Bernoulli), exprimez la distribution a posteriori en fonction de la distribution a priori et de la vraisemblance, puis explicitez la mise-à-jour des paramètres de la distribution conjuguée. *(1 point(s))*

Si on applique cela au problème de décisions séquentielles qu'est le bandit, on a à chaque instant une distribution a posteriori sur le paramètre de chaque bras. L'idée est alors de tirer pour chaque bras le paramètre associé selon la distribution a posterori, puis d'agir de façon gloutonne par rapport au résultat de ce tirage.

▷ **Q11** : Écrivez l'algorithme correspondant et expliquez informellement en quoi il permet de traiter le dilemme entre exploration et exploitation. *(1 point(s))*

Cette approche étant bayésienne, on peut utiliser naturellement l'information que l'on aurait a priori sur le problème.

▷ **Q12** : Comment tenir compte de cette information que l'on aurait a priori ? A contrario, quel est le meilleur choix si l'on n'a aucune information a priori (dans ce cas particulier, une réponse analytique est attendue) ? *(1 point(s))*

4 Renforcement et apprentissage du modèle (3 point(s))

On considère ici le cadre standard des processus décisionnels de Markov (PDM). On se donne un PDM $\{\mathcal{S}, \mathcal{A}, P, r, \gamma\}$, où l'espace d'état est fini et suffisamment petit pour permettre une représentation tabulaire de la fonction de valeur, et où l'espace d'action est également fini (et de petite taille). La fonction de récompense r est connue, mais la dynamique P du système est inconnue. Toutefois, on dispose d'une base de donnée de transitions de la forme

$$\mathcal{D} = \{(s_i, a_i, s'_i)_{1 \leq i \leq n}\},$$

où les états s_i et a_i ont été tirés uniformément, et les état s'_i selon la dynamique ($s'_i \sim P(.|s_i, a_i)$).

▷ **Q13** : On fixe un couple (s, a). Quelle type de loi suit $P(S'|S = s, A = a)$? Donnez le maximum de vraisemblance des paramètres de cette loi. *(1 point(s))*

▷ **Q14** : En déduire une façon d'estimer le modèle de transition. Comment l'utiliser pour estimer une politique optimale ? *(0,5 point(s))*

▷ **Q15** : Quel problème voyez vous avec cette approche (on pourra faire l'analogie avec une stratégie, pour un problème de bandit, qui consisterait à agir de façon gloutonne par rapport aux moyennes empiriques de chaque bras) ? *(0,5 point(s))*

Cette stratégie (d'estimation fréquentiste du modèle) pose donc le problème du dilemme entre exploration et exploitation (qu'elle ignore en fait). On souhaite affiner le modèle en interagissant avec le système.

▷ **Q16** : Est-il pertinent de simplement appliquer la politique optimale (respectivement au modèle estimé) pour collecter de nouvelles données ? *(0,5 point(s))*

▷ **Q17** : Une stratégie simple (appelée `Rmax`) consiste à modifier la récompense de la façon suivante. Soit $r_{\max}$ un majorant de la récompense, la nouvelle récompense r', paramétrée par un entier T, est définie par:

$$r'(s, a) = \begin{cases} r_{\max} & \text{si le couple } (s, a) \text{ a été expérimenté moins de } T \text{ fois} \\ r(s, a) & \text{sinon} \end{cases}.$$

Expliquez informellement en quoi cette stratégie est meilleure que celle de la question précédente. *(0,5 point(s))*