



3. (2 points) Une question ...  
(*A question ...*)

4. (1 point) Une question ...  
(*A question ...*)

5. (2 points) Une question ...  
(*A question ...*)

6. (1 point) Une question ...  
(*A question ...*)

7. (3 points) Supposez que vous décidez d'utiliser un perceptron multicouches pour apprendre à associer une classe ('+' ou '-') à la description d'un exemple.

- Que pouvez vous contrôler et modifier pour réaliser l'apprentissage ?  
(*What are the parameters that you can control in order to test several learning procedures?*)

- Supposons qu'après avoir expérimenté en modifiant extensivement les paramètres de contrôle, vous n'arrivez toujours qu'à une performance faible en généralisation, par exemple 65%, que faudrait-il en penser ? Que proposeriez-vous de faire ?

(*Suppose that after extensive experimentations varying all the possible control parameters you only get a poor prediction performance, for instance 65%, what would you conclude? What would you propose as a next step?*)

8. (1 points) Validation de l'apprentissage. Dites à quoi servent respectivement les *ensemble d'apprentissage*, *ensemble de test* et *ensemble de validation*.

(*What is the use of the training set, the test set and the validation set?*)

9. (2 points) Sur la validation croisée.

- Qu'est-ce que la validation croisée ? À quoi sert-elle ? Comment marche-t-elle ?  
(*What is cross-validation? What is it used for? How does it work?*)

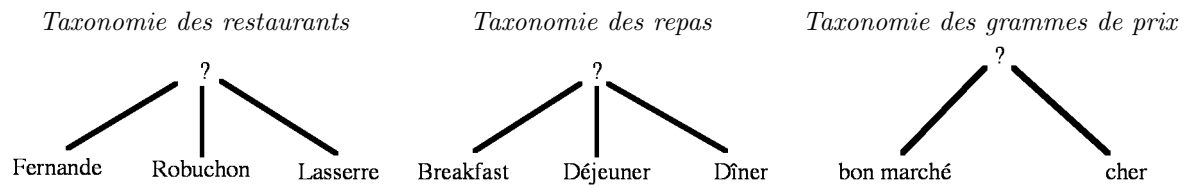
- Lorsque l'on utilise la validation croisée à  $n$  plis pour évaluer la performance en généralisation d'un apprentissage, on obtient à la fin  $n$  fonctions de décision différentes. Comment fait-on pour prédire la classe d'un nouvel exemple ?  
(*When using  $n$ -fold cross-validation in order to evaluate the generalization performance, one ends up with  $n$  decision functions. How can we proceed in order to make a prediction for a new unknown example?*)

## 2 Exercice : espace des versions (5 points)

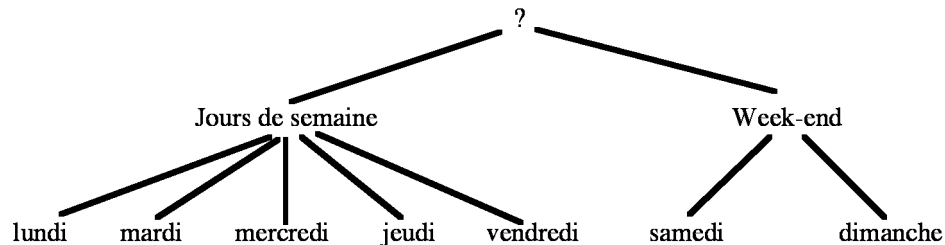
On suppose que les exemples fournis à l'algorithme sont décrits à l'aide de **quatre descripteurs** :

- le *nom du restaurant* : Robuchon, Fernande, Lasserre ;
- le *type de repas* : matin, déjeuner, dîner ;
- le *jour de la semaine* : lundi, mardi, mercredi, jeudi, vendredi, samedi, dimanche ;
- la *gamme de prix* : bon marché, cher.

On suppose de plus que les descripteurs sont organisés sous forme de hiérarchies dont le sommet (la racine) correspond au plus grand niveau de généralité, tandis que les feuilles correspondent au niveau le moins général (ainsi, dans l'arbre des descripteurs des feuilles ci-dessous, samedi est moins général que week-end, qui lui-même est moins général que semaine, par contre, il n'y a pas de lien de généralité entre samedi et lundi car ils appartiennent à des branches différentes de l'arbre).



*Taxonomie des jours de semaines :*



Soit la séquence d'exemples suivante :

- E1** : Fernande  $\wedge$  Breakfast  $\wedge$  Samedi  $\wedge$  Bon marché  $\rightarrow$  classe +  
**E2** : Robuchon  $\wedge$  Déjeuner  $\wedge$  Samedi  $\wedge$  Cher  $\rightarrow$  classe -  
**E3** : Fernande  $\wedge$  Déjeuner  $\wedge$  Dimanche  $\wedge$  Bon marché  $\rightarrow$  classe +  
**E4** : Lasserre  $\wedge$  Breakfast  $\wedge$  Lundi  $\wedge$  Cher  $\rightarrow$  classe -  
**E5** : Fernande  $\wedge$  Breakfast  $\wedge$  Dimanche  $\wedge$  Cher  $\rightarrow$  classe +

1. Spécifiez le S-set et le G-set après la prise en compte de chacun d'eux. (3 points)

2. Comment seront alors classés les exemples suivants et pourquoi ? (1 points)

- Fernande  $\wedge$  Dîner  $\wedge$  Dimanche  $\wedge$  Bon marché
- Fernande  $\wedge$  Déjeuner  $\wedge$  Vendredi  $\wedge$  Bon marché
- Robuchon  $\wedge$  Déjeuner  $\wedge$  Samedi  $\wedge$  Bon marché

3. Que signifierait l'épuisement de l'espace des versions après la prise en compte d'une séquence d'exemples et de contre-exemples ? (1 point)