

Documents multimédia : description et recherche automatique

M2P GI, examen du 22 avril 2015, 2 heures, documents autorisés.

Exercice 1. Modèles de langues - généralités

Expliquer quel est le rôle du lissage de probabilités, en se basant sur l'exemple deux dés d_1 et d_2 , pour lesquels les probabilités de tirage des 6 faces est le suivant :

Pour le dé d_1 :

$$P(1|d_1) = P(3|d_1) = P(5|d_1) = 0.3$$

$$P(4|d_1) = P(6|d_1) = 0.05$$

$$P(2|d_1) = 0$$

Pour le dé d_2 :

$$P(2|d_2) = P(4|d_2) = P(6|d_2) = 0.3$$

$$P(1|d_2) = 0$$

$$P(3|d_2) = P(5|d_2) = 0.05$$

Supposons que l'on observe la séquence $Q = \{1, 2, 3, 2\}$.

Question 1.1 : En utilisant les probabilités ci-dessus telles quelles, quelle est la probabilité pour le dé d_1 de générer la requête.

Question 1.2 : En utilisant les probabilités ci-dessus telles quelles, quelle est la probabilité pour le dé d_2 de générer la requête.

Question 1.3 : En faisant un parallèle avec les résultats obtenus sur les dés aux questions 1 et 2, indiquer pourquoi le lissage de probabilités, comme nous l'avons vu en cours, est très important pour la recherche d'images.

Exercice 2. Modèles de langues - lissages

Prenons comme exemple la description de deux images qui sont représentées par des occurrences de clusters :

	D1	D2
t	w_{D1}^t	w_{D2}^t
C1	1	0
C2	1	3
C3	3	0
C4	0	1
C5	1	0

Question 2.1

Calculer les modèles $P_{ML}(t|\theta_{D1})$ et $P_{ML}(t|\theta_{D2})$ pour ces deux documents. Les représenter sous forme d'un tableau similaire à celui fourni ci-dessus. (NOTE : Ces probabilités **ne sont pas** lissées).

Question 2.2

On réalise maintenant un lissage de Jelinek-Mercer avec $\lambda = 0.1$ pour les probabilités des documents.

Indiquer dans un tableau les probabilités lissées $P(t|\hat{\theta}_{D2})$ pour le document D2 sachant que :

- les probabilités $P(t|C)$ sont les suivantes :

P(t C)
0,1
0,4
0,3
0,1
0,1

- les probabilités sont les suivantes pour le document D1 :

	D1
t	$P(t \hat{\theta}_{D1})$
C1	0,160
C2	0,190
C3	0,480
C4	0,010
C5	0,160

Question 2.3

Supposons que l'on traite la requête Q suivante : {C2 C3 C3 C4}

Donner les similarités $P(Q|\hat{\theta}_{D1})$ et $P(Q|\hat{\theta}_{D2})$ entre Q et les documents D1 et D2 en utilisant le modèle de langue multinomial en se basant sur la formule du cours en transparent 14 (rappel w_t^Q est le nombre de fois où t apparaît dans Q) et en utilisant les probabilités calculées en question 3. En vous basant sur les calculs de similarités, indiquer si D1 répond mieux que D2 à la requête ou si c'est le contraire.

Question 2.4

Il est possible de penser à d'autres lissages. Une idée simple est de fixer une constante c, par exemple c=0.1, et :

1. Si n termes apparaissent dans un document D, on retire de chacune des probabilités P_{ML} de ces termes la valeur c/n. Ceci donne donc : $P'(t|\hat{\theta}_D) = P_{ML}(t|\theta_D) - \frac{c}{n}$
2. Si m termes n'apparaissent pas dans le document D, on fixe chacune des probabilités lissées à c/m. Ceci donne donc : $P'(t|\hat{\theta}_D) = \frac{c}{m}$

NOTE : m+n est alors égal au nombre de termes possibles (dans notre cas m+n=5)

Fournir le tableau de ces nouvelles probabilités P' pour le document D2 avec c=0.1, sachant que pour le document D1 nous avons :

	D1
t	$P'(t \hat{\theta}_{D1})$
C1	0,142
C2	0,142
C3	0,475
C4	0,100
C5	0,142

Question 2.5

Reprendre la question 3 avec les lissages de la question 4, et comparer les résultats obtenus avec ceux de la question 3.

Exercice 3. Descripteurs histogrammes



image 1

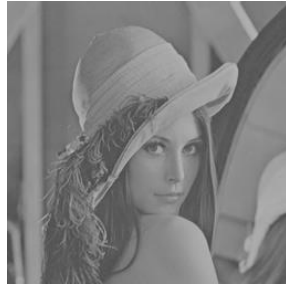


image 2



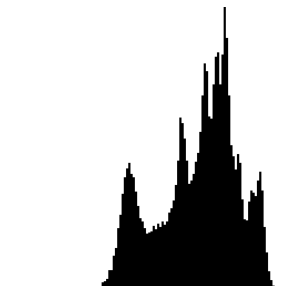
image 3



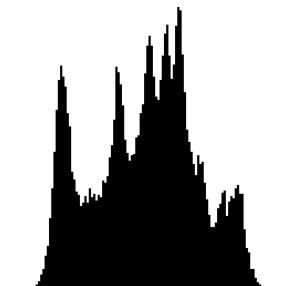
image 4



image 5



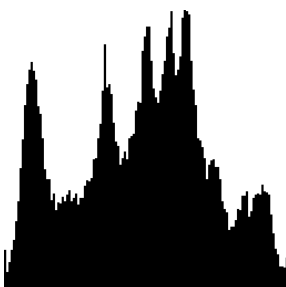
histogramme a



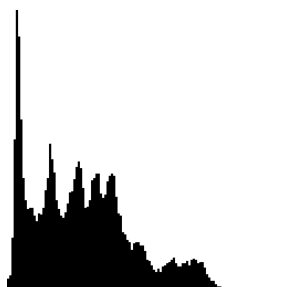
histogramme b



histogramme c



histogramme d



histogramme e

Question 3.1 : Associez les images ci-dessus avec leurs histogrammes, justifiez vos réponses. Les histogrammes sont monodimensionnels et portent sur l'intensité (luminance) de l'image.

Question 3.2 : Quelle est la dimension d'un descripteur d'image basé sur un histogramme de couleurs tridimensionnel dans l'espace LAB avec respectivement 8, 5 et 5 « bins » pour les composantes L, A et B ?

Question 3.3 : Même question si on calcule en plus cet histogramme par bloc avec un découpage des images suivant une grille bidimensionnelle avec 3 lignes et 4 colonnes ?

Question 3.4 : Parmi ces deux descripteurs (questions 3.2 et 3.3), lequel est sensible à une transformation par symétrie horizontale et lequel est robuste à une telle transformation ?

Question 3.5 : Quelles distances peuvent être utilisées pour évaluer la similarité de ces histogrammes dans le cadre d'une recherche par l'exemple ?

Question 3.6 : Ces distances sont-elles robustes par rapport à un changement de luminosité (brillance ou contraste) ?

Exercice 4. Classification par les k plus proches voisins et évaluation

On dispose d'une collection d'images exemples annotées selon la présence ou non d'un concept cible (une voiture par exemple). On dispose d'une autre collection d'images non annotées et pour lesquelles on voudrait produire des prédictions de présence ou non du même concept cible.

Les images de la collection sont nommées I_n , représentées par un descripteur x_n et annotée par l'étiquette a_n , avec $1 \leq n \leq N$. Les étiquettes a_n prennent la valeur 0 ou 1 selon que le concept cible est visible ou non dans l'image I_n . Les descripteurs x_n sont des vecteurs de nombre réels de dimension fixe d .

La prédiction est faite indépendamment pour chaque image de l'autre collection, on ne considèrera donc que le cas d'une seule image test I également représentée par un descripteur x de même type. On supposera que la distance euclidienne est adaptée pour la comparaison de descripteurs de ce type.

Question 4.1 : Donnez un algorithme pour prédire l'étiquette a correspondant à la visibilité du concept cible dans l'image I selon la méthode des k plus proches voisins (k nearest neighbors) avec $k = 5$ et une règle de décision basée sur une vote majoritaire.

Question 4.2 : La décision basée sur un vote majoritaire fournit seulement une réponse binaire qui ne permet pas de trier les images non annotées de la plus probablement positive à la plus probablement négative. Proposez une variante de l'algorithme précédent prenant en compte la distance de l'image à annoter à ses plus proches voisins trouvés pour produire un score de confiance réel permettant un tel classement. Plusieurs solutions sont possibles, une seule suffit.

Question 4.3 : Quelles mesures d'évaluation peut-on utiliser pour évaluer un système qui fournit une liste non ordonnée (un ensemble) de résultats ou seulement une valeur booléenne indiquant la pertinence ou non pertinence de chaque résultat (comme dans la question 4.1) ?

Question 4.4 : Quelles mesures d'évaluation peut-on utiliser pour évaluer un système qui fournit une liste ordonnée de résultats ou un score de confiance permettant un tri des résultats du plus probablement pertinent au moins probablement pertinent (comme dans la question 4.2) ?