



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Departamento de sistemas informáticos y
computación

Predicción Estructurada Estadística

Cuestionario Parte 2

Máster Universitario en Inteligencia Artificial, Reconocimiento de Formas
e Imagen Digital

Autor

Francisco Javier Gil-Terrón Rodríguez

2021 - 2022

Tabla de contenidos

1.	Cuestión 1	3
2.	Cuestión 2	3
3.	Cuestión 3	4
4.	Cuestión 7	5

1. Cuestión 1

Briefly explain the differences between Classification and Structured Output Prediction. Cite two application examples each paradigm.

Mientras que en un problema de clasificación se trata de asignar a una muestra dada una etiqueta correcta de entre un conjunto dado de clases, siendo cada resultado independiente de cualquier otro producido por otra muestra, en el caso de la predicción estructurada el resultado es un conjunto que puede ser dependiente mutuamente unos de otros y por tanto deben ser tenidos en consideración, además de que el espacio del problema es potencialmente infinito.

Algunos ejemplos de clasificación podrían ser análisis de sentimientos en textos, donde a partir de un texto se pretende extraer si las emociones que transmiten son positivas, negativas o neutras; o la detección de tumores en imagen médica, donde el objetivo es clasificar una imagen en función de si tiene tumor o no.

Por otro lado, en cuanto a predicción estructurada, algunos casos de aplicación serían la generación automática de subtítulos a partir de un audio, o la generación de un texto determinado a partir de otro (ya sea un resumen, síntesis o ampliación).

2. Cuestión 2

Justify why the naïve Bayes decomposition of Eq.(5) is adequate for karyotype recognition problem.

En la fórmula siguiente se refleja este problema donde se deben etiquetar imágenes de cromosomas humanos teñidos de un conjunto de 22 etiquetas (los distintos cromosomas), de manera que x serán las muestras de imágenes y h las etiquetas de los cromosomas.

$$P(x | h) = P(x_1, \dots, x_{22} | h_1, \dots, h_{22}) \approx \prod_{i=1}^{22} P(x_i | h_i)$$

La asunción de Naive Bayes, pese a que asume que las características de las muestras son independientes entre sí, puede llegar a ser útil incluso en aquellos casos donde las características no son independientes ya que permite reducir la complejidad y, por tanto, el tiempo de cómputo, haciendo que el algoritmo sea más rápido, ya que de esta manera no se evitará estimar una gran cantidad de parámetros. Con este en mente, la ecuación anterior podría aproximarse a la siguiente:

$$\approx P(x_1|h_1)P(x_2|h_2) \dots P(x_{22}|h_{22})$$

Para ello se deberá tener en consideración que los cromosomas son independientes entre sí, ya que la forma de un cromosoma no depende de la de otro cromosoma, y que las etiquetas también lo son, ya que la representación de un cromosoma en concreto es distinta a la cualquier otro.

3. Cuestión 3

Briefly explain all the steps and assumptions needed to derive Eq.(9) from Eq.(7).

La ecuación que se mostrará a continuación establece que dada una muestra x , una historia h' y retroalimentación f , la hipótesis óptima sigue la expresión:

$$\hat{h} = \operatorname{argmax}_h P(h|x, h', f)$$

En un entorno de retroalimentación determinista, se puede definir una función de decodificación $d = d(f)$, con lo que no es necesario tener un modelo para reconocer la retroalimentación. Sustituyéndolo en la ecuación anterior y aplicando la regla conjunta se obtendrían las siguientes:

$$\hat{h} = \operatorname{argmax}_h P(h|x, h', d) = \operatorname{argmax}_h \frac{P(h, x, h', d)}{P(x, h', d)}$$

Además, podemos ignorar el denominador, ya que no depende del parámetro h que se está maximizando. Tras esto, aplicando la regla de la cadena se obtiene:

$$\hat{h} = \operatorname{argmax}_h P(h, x, h', d) = \operatorname{argmax}_h P(h')P(d|h')P(h|h', d)P(x|h)$$

Finalmente, se puede limpiar la ecuación eliminando los elementos no dependientes de h , es decir, aquellos elementos que serán constantes en la maximización, obteniendo la ecuación objetivo.

$$\hat{h} = \operatorname{argmax}_h P(h')P(d|h')P(h|h', d)P(x|h) = \operatorname{argmax}_h P(x|h)P(h|h', d)$$

4. Cuestión 7

Briefly explain under which conditions the solution given by Eq.(22-23) may be optimal. Do the same conditions hold for the optimality of the solution given by Eq.(20-21)? Why? Use the karyotyping example to illustrate your (otherwise general) responses.

En primer lugar, las ecuaciones 22 y 23 de las diapositivas serían óptimas en el caso de que las n codificaciones más probables tuvieran la misma longitud que la talla del problema. Trasladado al problema de los cariotipos, esto significa un tamaño de 22 (el número de cromosomas). En caso de considerar que el mismo número de decodificaciones es igual a la longitud de la talla del problema, se obtiene retroalimentación para corregir todos los errores y calcular la hipótesis óptima, en otras palabras, se tiene toda la información posible para corregir los errores.

En cuanto a las ecuaciones 20 y 21 de las diapositivas, considerando las mismas condiciones expuestas en el punto anterior, no serán óptimas, ya que en este caso obtienen primero una decodificación óptima para la retroalimentación y, tras esto, obtienen una hipótesis óptima en base a esa decodificación. Bajo esta consideración, pueden generarse distintas combinaciones no óptimas, siendo el número de estas combinaciones mayor al de la talla del problema, por lo que acotando que n fuera igual a dicha talla, no se puede garantizar que se obtenga una hipótesis óptima, ya que podría existir una combinación no óptima que permitiera obtener una hipótesis mejor.