Profesor: Gonzalo Ríos
Auxiliar: Miguel Romero
Fecha: 30 de Abril

Examen

1 Pregunta 1

El cachipún es un juego entre 2 jugadores, en donde en cada turno, cada jugador elige una de las tres opciones: "piedra", "papel", "tijera". La siguiente matriz muestra la resolución de cada turno:

	piedra	papel	tijera
piedra	empate	papel	piedra
papel	papel	empate	tijera
tijera	piedra	tijera	empate

el fin de esta pregunta es construir un modelo "inteligente" para jugar cachipún, usando probabilidades para este fin.

- Una N-partida de cachipún es una secuencia de N-tuplas de la forma $\{(J_i, O_i)\}_{i=1}^N$, en donde J_i indica la elección del jugador en el i-ésimo turno, y O_i indica la elección del oponente en el i-ésimo turno.
- La elección del jugador en el turno k solo dependerá de la elección del oponente en el turno $k - 1$ (no del número del turno), y esta será probabilista. Por ejemplo, el modelo podrá estimar la probabilidad de ganar el turno si saca piedra, dado que el turno pasado el oponente sacó tijera.
- El modelo deberá partir de la condición apriori equiprobable, e ir aprendiendo la estrategia del oponente turno a turno.
- Para esto, el modelo deberá estimar

$$\Pr(\text{El oponente saque } C \text{ en este turno} \mid \text{El oponente sacó } X \text{ en el turno pasado}) (*)$$

1. Indique que tipo de tarea de Minería de Datos es este problema. (0.5 pto)
2. Usando el Teorema de Bayes, encuentre una expresión equivalente para (*) (0.5 pto)
3. Dada la N-partida $\{(J_i, O_i)\}_{i=1}^N$, usando el estimador de máxima verosimilitud y la expresión anterior, dé una fórmula para calcular (*). ¿Cómo cambia esta fórmula si hubiera estimado directamente (*) con el EMV? (1.5 pto)
4. Dada la secuencia del oponente { piedra,papel,tijera,piedra,papel,tijera,piedra,papel,tijera }, haga la predicción de elección del oponente en el próximo turno (con sus respectivas probabilidades). Si su oponente saca piedra en el próximo turno, que puede decir de su modelo? (1.5 pto)
5. Detalle los procedimientos de su modelo en una partida real. ¿Su modelo es lazy o eager?. ¿Qué optimización podría hacer a su modelo? (1 pto)
6. ¿Cómo cambia su algoritmo si en vez de 3 opciones, cada jugador tiene M opciones? Estime la cantidad de parámetros del modelo. (1 pto)

2 Pregunta 2

1. Defina Validación Cruzada y Holdout.
2. Explique el funcionamiento del algoritmo de K Medias para clustering.
3. Dada una serie de tiempo $\{Y_i\}_{i=1}^T$, plantee el problema de optimización, usando error relativo, para construir un modelo autorregresivo de orden p .
4. Dado el ordenamiento de variables F,E,S,G, y sabiendo que se cumple $\Pr(e|f) = \Pr(e)$, $\Pr(s|f,e) = \Pr(s)$, $\Pr(g|f,e,s) = \Pr(g|f)$, construya la red bayesiana óptima.
5. Diga los pro y los contra de las Redes Neuronales.
6. Explique el Algoritmo de Hunt, incluyendo el criterio de selección del mejor split.