



Comenzado el

 **miércoles, 11 de septiembre de 2024, 20:58**

Estado

 **Finalizado**

Finalizado en

 **miércoles, 11 de septiembre de 2024, 21:11**

Tiempo empleado

 **12 minutos 56 segundos**

Puntos

 **4,10/6,00**

Calificación

 **3,42 de 5,00 (68,33%)**

Pregunta 1

Correcta

Se puntúa 1,00 sobre 1,00

Si vamos a entrenar una red neuronal tradicional para clasificación, cuáles de los siguientes algoritmos de optimización funcionarían para la búsqueda de los pesos de la red?

Seleccione una:

☒ a. Todos los mencionados aquí. ✓

☐ b. Stochastic gradient descent.

☐ c. Batch gradient descent.

☐ d. Mini batch gradient descent.

**Respuesta correcta**

La respuesta correcta es: Todos los mencionados aquí.

Pregunta 2

Parcialmente correcta

Se puntúa 1,50 sobre 2,00

En el contexto del gradient descent, emparejar las afirmaciones con las opciones **más adecuadas**:

Es mas susceptible a mínimos locales.

Stochastic gradient descent.



Presenta muchísima varianza en la búsqueda de la convergencia de la función de costo.

Stochastic gradient descent.



Permite actualizar los parámetros varias veces en una época.

Stochastic y mini batch gradient descent.



Se puede demorar muchas épocas en llegar a convergencia de la función de costo.

Batch gradient descent.



Respuesta parcialmente correcta.

Ha seleccionado correctamente 3.

La respuesta correcta es: Es mas susceptible a mínimos locales. → Batch gradient descent., Presenta muchísima varianza en la búsqueda de la convergencia de la función de costo. → Stochastic gradient descent., Permite actualizar los parámetros varias veces en una época. → Stochastic y mini batch gradient descent., Se puede demorar muchas épocas en llegar a convergencia de la función de costo. → Batch gradient descent.

Pregunta 3

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes opciones no es un hiperparámetro de los algoritmos de gradient descent?

Seleccione una:

- ☐ a. Learning rate.
- ☐ b. Ninguno de los mencionados aquí.
- ☒ c. Nivel de sesgo. ✓
- ☐ d. Número de épocas de entrenamiento.
- ☐ e. Tamaño del batch.

**Respuesta correcta**

La respuesta correcta es: Nivel de sesgo.

Pregunta 4

Parcialmente correcta

Se puntúa 0,60 sobre 1,00

Empareje las afirmaciones con las opciones **más adecuadas**, en el contexto de un proceso de optimización por descenso de gradiente.

Un tamaño de de batch de entrenamiento menor siempre es mejor.

FALSO, no tiene ninguna incidencia. ▾

✗

Un mayor de tamaño del batch de **evaluación** permite llegar más rápidamente a convergencia.

VERDADERO, aunque hay que tener cuidado con la divergencia. ▾

✗

Un mayor learning rate permite aprender más rápido los parámetros de la red.

VERDADERO, aunque hay que tener cuidado con la divergencia. ▾

✓

Un menor learning rate permite aprender de una manera menos caótica.

VERDADERO, aunque es mas susceptible a óptimos locales. ▾

✓

Un mayor de tamaño del batch de **entrenamiento** presenta menos varianza en el entrenamiento.

VERDADERO, aunque hay que tener cuidado con la divergencia. ▾

✗



Respuesta parcialmente correcta.

Ha seleccionado correctamente 2.

La respuesta correcta es:

Un tamaño de de batch de entrenamiento menor siempre es mejor. → FALSO, se puede quedar atascado en un mínimo local.,

Un mayor de tamaño del batch de evaluación permite llegar más rápidamente a convergencia. → FALSO, no tiene ninguna incidencia.,

Un mayor learning rate permite aprender más rápido los parámetros de la red. → VERDADERO, aunque hay que tener cuidado con la divergencia.,

Un menor learning rate permite aprender de una manera menos caótica. → VERDADERO, aunque es mas susceptible a óptimos locales.,

Un mayor de tamaño del batch de entrenamiento presenta menos varianza en el entrenamiento. → FALSO, se puede quedar atascado en mínimos locales.

Comentario:

Pregunta 5

Incorrecta

Se puntúa 0,00 sobre 1,00

Con respecto a la técnica de "**Early Stopping**", escoja la(s) opción(es) **más adecuada(s)**:

Seleccione una o más de una:

- ☐ a. Solo se puede aplicar con Batch gradient descent.
- ☒ b. Independientemente de si es con Batch, Mini batch o Stochastic gradient descent, se parará el entrenamiento apenas la función de costo deja de decrecer. ×
- ☐ c. Se puede utilizar para establecer el mejor learning rate.
- ☐ d. Permite establecer el número de épocas adecuado dados los demás hiperparámetros.
- ☐ e. Es una manera de luchar contra el overfitting.



Respuesta incorrecta.

Las respuestas correctas son: Es una manera de luchar contra el overfitting,,
Permite establecer el número de épocas adecuado dados los demás hiperparámetros.

[Finalizar revisión](#)

🌐 [Intu Universidad Icesi](#)

☎ Teléfono: 5552334 ext 4188

✉ icesivirtual.ti@listas.icesi.edu.co



👤 **Adaptación:** [Icesi Virtual](#) | [Cristian Vásquez](#)

Copyright © 2024 Universidad Icesi

[Contactar con el soporte del sitio](#)

[Resumen de retención de datos](#)

[Descargar la app para dispositivos móviles](#)

