

SegundoExamen Parcial

Sobre la Google Net, cual de las siguientes afirmaciones no es verdadera 2/2

- ☐ Utilizó bloques que encapsulan filtros de 1x1, 3x3 y 5x5 (Inception 1)
- ☐ La principal desventaja que tuvo fue su topología heterogenia
- ☐ El principal objetivo de la Arquitectura fue la de mejorar el accuracy y, a la par, reducir el costo computacional
- ☒ NA

Cual de la siguientes afirmaciones es falsa 4/4

- ☐ Para lidiar con el problema de pérdida de información generada por kernels grandes, se amplió la red en anchura utilizando bancos de filtros más pequeños que podian ser calculados en paralelo
- ☒ Una convolución de NxN es menos costosa que una convolución de 1xn y otra de nx1
- ☐ Uno de los métodos para lidiar con el desvanecimiento de la gradiente es utilizando ReLu como función de activación
- ☐ El Dropout es un método de regularización



Sobre las RNN Bidireccionales marque la respuesta falsa

2/2

- ☐ Son utilizadas para explorar el contexto dependiendo tanto del pasado como del futuro.
- ☐ Las salidas del estado de las celdas forward no tiene conexión con las entradas de las celdas backward
- ☒ No requieren conocer el inicio y el final de una secuencia de entrada.
- ☐ Pueden ser entrenadas mediante BPTT pero es ligeramente más complicado que en el caso de las RNN no bidireccionales.
- ☐ Otro:

¿Cuál de las siguientes funciones genera el problema de saturación de la gradiente 1). Sigmodie, 2). Tanh. 3). Relu.

1/1

- ☐ Sólo 1
- ☐ Sólo 2
- ☐ Sólo 3
- ☒ 1 y 2
- ☐ 2 y 3
- ☐ 1,2 y 3



En el caso de la regularización L2, respecto a la regresión no lineal, se puede 1/1 afirmar que: 1) Si el parámetro regularizador es muy grande es posible llegar a un underfitting. 2) Si el parámetro regularizado es muy pequeño se puede llegar a un overfitting

- ☐ Solo 1
- ☐ Solo 2
- ☒ 1 y 2
- ☐ Ninguno

Digamos que está utilizando la función de activación X en capas ocultas de la 1/1 red neuronal. En una neurona particular para cualquier entrada dada, obtiene la salida como "-0.0001". ¿Cuál de las siguientes funciones de activación podría representar X?

- ☐ Relu
- ☒ tanh
- ☐ Sigmoide
- ☐ NA

Lasso puede ser interpretado como una regresión lineal donde

1/1

- ☒ Los pesos son regularizados con la norma L1
- ☐ Los pesos son regularizados con la norma L2
- ☐ El parámetro regularizador debe ser un valor alto
- ☐ Puede usar L1 o L2



Sobre la Alex Net, cual de las siguientes afirmaciones no es verdadera

2/2

- ☐ Utilizó filtros grandes (11x11 y 5x5)
- ☐ Aplicó submuestreo superpuesto para mejorar la generalización y reducir el overfitting
- ☒ Considerada como la primera Deep CNN
- ☐ NA

Sobre los kernel

1/1

- ☐ Puede ser aplicado a todo algoritmo de clasificación
- ☐ Es común utilizarlo en reducción de la dimensionalidad
- ☒ explota el hecho que los pesos pueden ser escritos como una combinación lineal de los datos de entrada.
- ☐ Ninguna

En las redes neuronales, las funciones de activación no lineal como sigmoide, tanh y ReLU

- ☐ Acelerar el cálculo de gradiente en retropropagación, en comparación con la unidad lineal
- ☐ Se aplican solo a las unidades de salida
- ☒ Ayudar a aprender los límites de decisión no lineales
- ☐ Los valores de salida son, siempre, entre 0 y 1



¿Cuál de las siguientes afirmaciones es verdadera sobre la Gradient Decent (GD) y Stochastic Gradient Decent (SGD)? 1/1

1). En GD y SGD, actualiza un conjunto de parámetros de manera iterativa para minimizar la función de error.

2). En SGD, debe ejecutar todas las muestras en su conjunto de entrenamiento para una única actualización de un parámetro en cada iteración.

3). En GD, usas los datos completos o un subconjunto de datos de entrenamiento para actualizar un parámetro en cada iteración.

- ☒ Sólo 1
- ☐ Sólo 2
- ☐ Sólo 3
- ☐ 1 y 2
- ☐ 2 y 3
- ☐ 1,2 y 3

Sobre las RNN Marque la o las respuestas correctas

2/2

- ☒ Son un tipo de modelo de aprendizaje de máquinas hechas de neuronas artificiales con uno o más bucles de retroalimentación
- ☒ Uno de los problemas con estas redes es el hecho de poder reconocer correlaciones entre conjuntos de datos que están distantes en el tiempo.
- ☐ Uno de los métodos que se utiliza en el proceso de entrenamiento es la conocida Gradiente Desendiente.
- ☒ La gradiente, en el proceso de entrenamiento, podría explotar en secuencias largas cuandos los pesos de la red son grandes.



Sobre la ZFnet, cual de las siguientes afirmaciones no es verdadera

2/2

- ☐ Fue Desarrollada para vizualizar la eficiencia de la redes de manera cuantitativa
- ☒ Su arquitectura fue modular en capas y llegó a tener 19 capas profundas.
- ☒ Remplazó los filtros de (11x11 y 5x5) en pilas de filtros de 3x3
- ☐ NA

Si entrenamos una RNN para estimar la distribución condicional del siguiente elemento de una secuencia y_t dado su pasado $y_1, \dots, y_{(t-1)}$, la función objetivo que nos conviene usar es: (Marque la respuesta correcta) 2/2

- ☐ SoftMax
- ☒ Log- Máximo Likelihood
- ☐ RELU
- ☐ Sigmoides
- ☐ Otro:

En un SVM, si el parámetro C , de la función de pérdida, tiende a cero podemos afirmar que: (Seleccione todas las alternativas correctas)

1/1

- ☒ Se incrementa el margen de decisión
- ☐ Decrementa el margen de decisión
- ☒ Se permiten algunos elementos mal clasificados
- ☐ No se permite ningún elemento mal clasificado
- ☐ Otro:



Es un método eficiente para extraer características similares en múltiples localizaciones con pocos parámetros 2/2

- ☐ Pooling
- ☒ Convolución
- ☐ Dropout
- ☐ Regularización

Sobre Mini-Batch Gradient Descendent (Marque la respuesta correcta) 1/1

- ☒ Computa la gradiente en batchs, donde cada batch tiene más de un ejemplo de entrenamiento.
- ☐ No es utilizado en las RNN
- ☐ Son difíciles de paralelizar
- ☐ Su performance suele ser peor que la GD y la SGD

Considere una capa de una red neuronal convolucional (CNN) para imágenes en escala de grises, conectada a una siguiente capa. ¿Qué tipo de capa tiene la menor cantidad de parámetros para aprender durante el entrenamiento? 0/1

- ☐ Una capa de convolución de 10 filtros de 3x3
- ☐ Una capa de max-pooling que reduce una imagen de 10x10 a 5x5
- ☒ Una capa convolucional de 8 filtros de 5x5
- ☐ Una full conectada de 20 unidades hidden conectada a 4 unidades de salida



Sobre las redes GRU marque la o las respuestas correctas

1/1

- ☒ Permiten disminuir el número de parámetros de las LSTM.
- ☒ A diferencia de las LSTM que tienen 3 compuertas, éstas tienen sólo 2.
- ☒ Permiten minimizar el desvanecimiento y la explosión de la gradiente.
- ☒ Son una variante de las LSTM.

Sobre las LSTM. ¿Cuál de las siguientes afirmaciones no es verdadera?

1/1

- ☐ Preservan las señales y propagan el error por mucho más tiempo que las RNN ordinarias
- ☐ Cada celda está hecha de una puerta de entrada, una puerta de salida y una puerta de olvido
- ☐ Las puertas LSTM pueden evitar que el resto de la red modifique el contenido de las celdas de memoria durante varios pasos de tiempo
- ☒ Ninguna de las anteriores

Sobre el problema de la explosión de la gradiente (marque la o las sentencias verdaderas) 2/2

- ☒ Es uno de los mayores problemas que se da durante el entrenamiento con BPTT
- ☒ Este problema ocurre cuando los pesos de la red tienen valores altos
- ☒ Uno de los métodos para abordar este problema es mediante la SGD
- ☒ Un ratio de aprendizaje adaptativo puede acelerar la convergencia hacia un mínimo local



Sobre la función de pérdida de la regresión logística podemos afirmar. 1) Para 1/1 un ejemplo que se sabe pertenece a la clase 1 y el clasificador le asigna una probabilidad baja, entonces el error es alto. 2) Cuanto mayor es la probabilidad que retorna el modelo, mejor es la clasificación

- ☒ Sólo la 1
- ☐ Sólo la 2
- ☐ La 1 y la 2
- ☐ Ninguna

Sobre ResNet, cual de las siguientes afirmaciones es verdadera

2/2

- ☒ Sugirió que el uso de funciones residuales son simples de optimizar y además puede mejorar el accuracy en redes más profundas.
- ☒ Introdujo conexiones entre capas para permitir conectividad entre estas
- ☒ Utiliza el concepto de función residual para lidiar con el problema del desvanecimiento de la gradiente
- ☐ NA

El campo receptivo se refiere a

2/2

- ☐ La matriz inicial de entrada de la CNN
- ☐ Los filtros en la CNN
- ☐ El conjunto de pesos de una capa C_k de la CNN
- ☒ Pequeños slides del mismo tamaño de los filtros dentro de la imagen

Este formulario se creó en Universidad Nacional de San Agustín.

Google Formularios

