

Técnicas Avanzadas de Data Mining y Sistemas Inteligentes

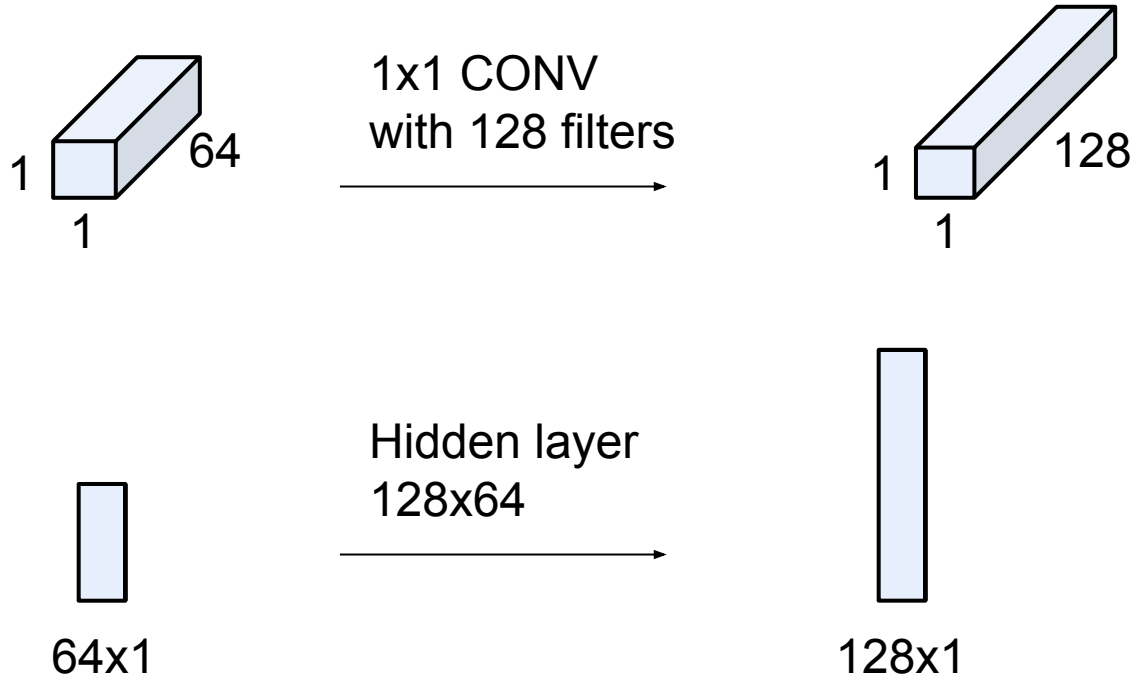
Maestría en Informática
Escuela de Posgrado
Pontificia Universidad Católica del Perú

2018-2

Review

1x1 convolution layers

Ej: 128 filtros de 1x1 en un input de 1x1x64



Transfer Learning - entrenamiento

Ej.
lr = 0.01



Modelo Pre-entrenado
(Solo las capas
convolucionales)

Features

Ej.
lr = 0.1

Predicción

Modelo de
Clasificación
(Capas densas)

Para un mejor resultado, se pueden entrenar también los pesos de la red pre-entrenada (es recomendable hacerlo con un learning rate menor: [Learning-Rate-Multipliers-in-Keras](#)).

Embeddings

¿Cómo ingresamos variables categóricas en una red neuronal?

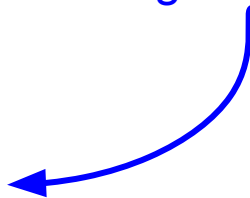
Variable categórica
cat 1
cat 2
cat 1
cat 3
cat 4
cat 2
cat 1
cat 4

Embeddings!

¿Cómo ingresamos variables categóricas en una red neuronal?

Variable categórica	ID
cat 1	0
cat 2	1
cat 1	0
cat 3	2
cat 4	3
cat 2	1
cat 1	0
cat 4	3


1. Asignamos números enteros como IDs a las categorías.



¿Cómo ingresamos variables categóricas en una red neuronal?

Variable categórica	ID
cat 1	0
cat 2	1
cat 1	0
cat 3	2
cat 4	3
cat 2	1
cat 1	0
cat 4	3

2. Obtenemos la lista de IDs únicos.

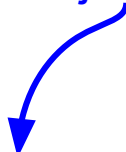


ID
0
1
2
3

¿Cómo ingresamos variables categóricas en una red neuronal?

Variable categórica	ID
cat 1	0
cat 2	1
cat 1	0
cat 3	2
cat 4	3
cat 2	1
cat 1	0
cat 4	3

3. Definimos la cantidad de variables que van a describir cada categoría. Por ejemplo: 5



ID	var1	var2	var3	var4	var5
0					
1					
2					
3					

¿Cómo ingresamos variables categóricas en una red neuronal?

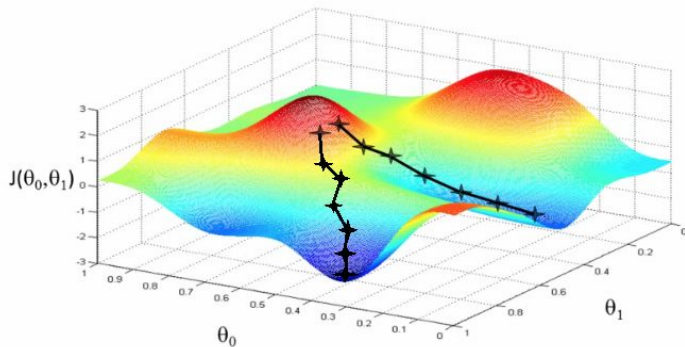
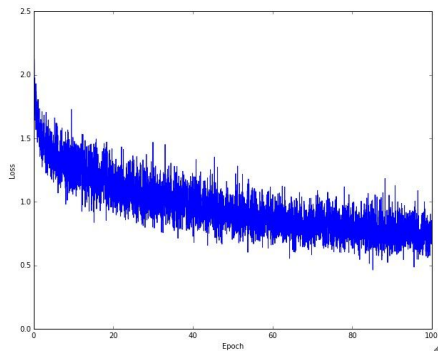
Variable categórica	ID
cat 1	0
cat 2	1
cat 1	0
cat 3	2
cat 4	3
cat 2	1
cat 1	0
cat 4	3

3. Como todo peso que se define en una red neuronal, se inicializa con números aleatorios.



ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

4. Y luego se optimizan los valores, para minimizar una función de pérdida..



ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

[illegible]

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

Usamos la matriz de embeddings como un lookup table

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

	var1	var2	var3	var4	var5
▶					

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

Usamos la matriz de embeddings como un lookup table

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

var1	var2	var3	var4	var5
0.04	0.78	0.09	0.63	0.01
0.04	0.78	0.09	0.63	0.01
0.04	0.78	0.09	0.63	0.01

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

Usamos la matriz de embeddings como un lookup table

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

var1	var2	var3	var4	var5
0.04	0.78	0.09	0.63	0.01
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

Usamos la matriz de embeddings como un lookup table

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

var1	var2	var3	var4	var5
0.04	0.78	0.09	0.63	0.01
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01
0.01	0.65	0.87	0.61	0.81
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01

¿Qué operación podemos usar para transformar los IDs en las variables latentes que definimos?

Usamos la matriz de embeddings como un lookup table

ID
0
1
0
2
3
1
0
3

ID	var1	var2	var3	var4	var5
0	0.04	0.78	0.09	0.63	0.01
1	0.94	0.43	0.25	0.19	0.41
2	0.01	0.65	0.87	0.61	0.81
3	0.83	0.27	0.84	0.28	0.14

var1	var2	var3	var4	var5
0.04	0.78	0.09	0.63	0.01
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01
0.01	0.65	0.87	0.61	0.81
0.83	0.27	0.84	0.28	0.14
0.94	0.43	0.25	0.19	0.41
0.04	0.78	0.09	0.63	0.01
0.83	0.27	0.84	0.28	0.14

Code time

<https://www.kaggle.com/renatoh/sound-classification-pucp>

<https://www.kaggle.com/renatoh/facial-keypoints-detection-pucp>

Extra:

<https://deepsense.ai/deep-learning-right-whale-recognition-kaggle/>

