



Trabajo Práctico N°2

Entrega Final

GRUPO N° 14

Apellido y Nombres	Dirección de E-Mail
Noe Tabarez	ctabarez@est.frba.utn.edu.ar
Maria Antonella Miozzo	mariamiozzo@est.frba.utn.edu.ar
Martin Fleichman	mfleichman@est.frba.tun.edu.ar
Ramiro Calvo	rcalvo@est.frba.utn.edu.ar
Alfredo Santiago Bose	abose@est.frba.utn.edu.ar

Fecha de Presentación	25/10/2020
------------------------------	------------

Universidad Tecnológica Nacional
Facultad Regional Buenos Aires



UTN.BA

FACULTAD
REGIONAL
BUENOS AIRES

Sistemas Inteligentes

Alfredo Santiago Bose, Ramiro Calvo, Martin Fleischman, Maria Antonella Miozzo, Noe Tabarez

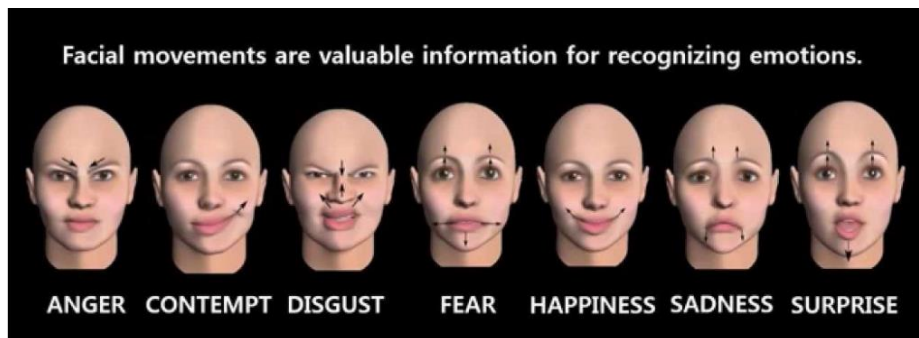
Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
Buenos Aires. Argentina

Resumen: Comprender las particularidades de la Implementación de un Sistema Inteligente mediante el uso de una red neuronal artificial (RNA)

Palabras Claves: CNN, Reconocimiento Facial, Estado de Animo, Procesamiento de Imagenes

1. Introducción

Cuántas veces nos ha pasado de que una persona de nuestro vínculo social nos ha dicho o preguntado, - *Te veo triste estás bien?*, - *Estas contenta qué ganaste?* y un sinfín de observaciones en base a nuestras expresiones faciales. Comentarios que creemos que se dan en base a una apreciación subjetiva y personal de un individuo que nos ve, y determina o relaciona que una sonrisa es sinónimo de felicidad y un gesto antónimo a la felicidad se relaciona con tristeza.



A raíz de esto buscamos resolver esta cuestión de una manera objetiva, que nos permita determinar el estado de ánimo de una persona a partir del análisis de sus expresiones faciales.

Esto nos brindaría un sinfín de beneficios a la hora de ofrecer un producto, en base a que tan aceptado es por el cliente, una comida si es rica o no, entre otras opciones que podríamos obtener si incorporamos al cliente como punto intermedio.

Una correcta solución de Inteligencia Artificial nos aportaría un sinfín de aplicación viables para esto, lo cual permitiría una nueva herramienta para determinar el grado de satisfacción de un cliente a la hora de adquirir, obtener algún producto.

2. Elementos de Trabajo y Metodología

Para poder llevar adelante esta problemática compleja, es necesario contar con una base de conocimiento/información importante para poder resolver esta cuestión. Para ello utilizaremos el dataset² brindado en una competencia de ciencias de datos¹, en donde el objetivo a resolver era similar al planteado por nosotros. Este dataset cuenta con 35887 registros de datos para llevar adelante nuestra problemática. Los datos que disponemos son:

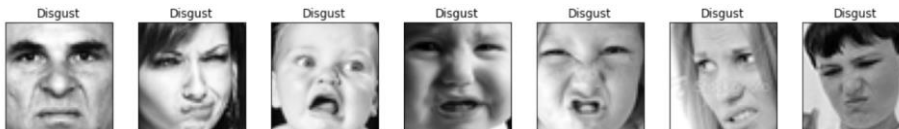
- La expresión Facial
- La imagen del individuo en pixels de 48x48

A continuación detallamos las expresiones faciales que se utilizan para diseñar la solución:

- Angry (Class 0)



- Disgust (Class 1)



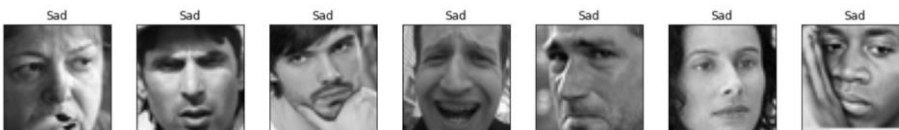
- Fear (Class 2)



- Happy (Class 3)



- Sad (Class 4)



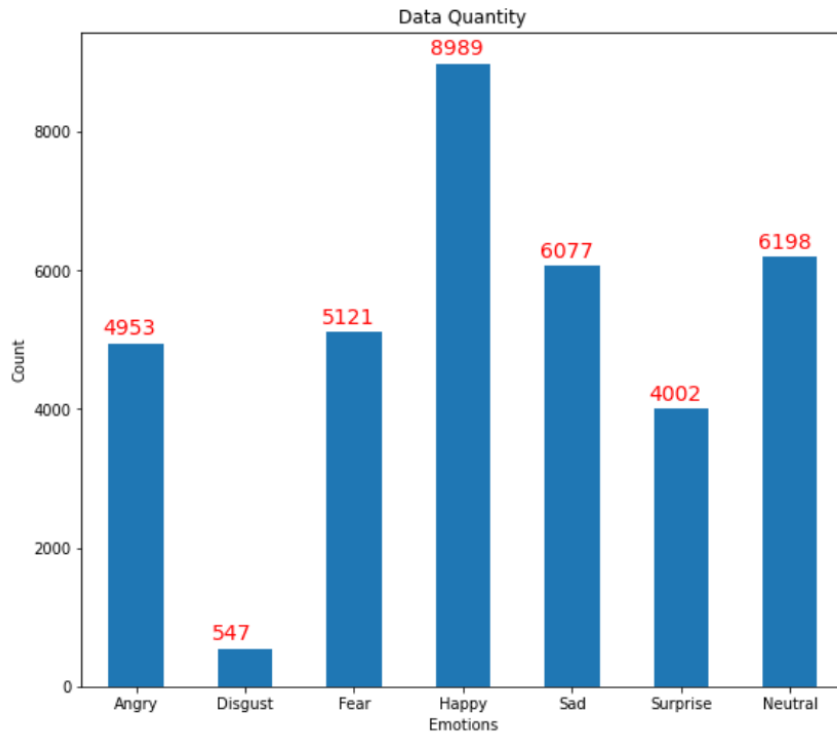
- Surprise (Class 5)



- Neutral (Class 6)



A modo de resumen, contamos con la siguiente cantidad de datos:



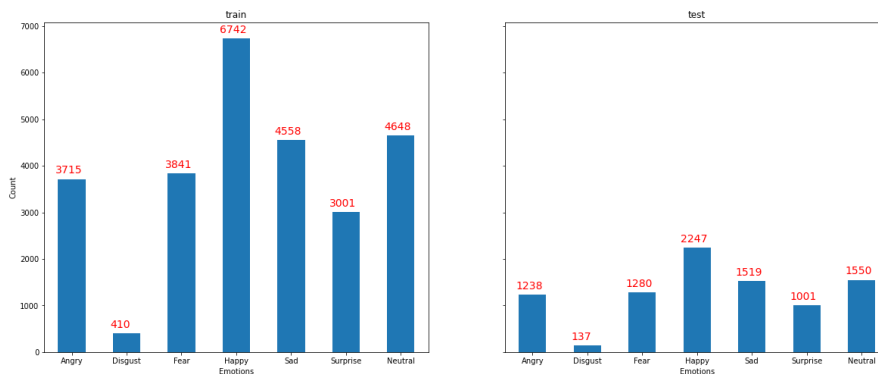
Para llevar adelante con nuestro objetivo, implementaremos una Red Neuronal Artificial MultiPerceptrón BackPropagation, ya que consideramos que es la que más se ajusta con nuestra solución que implícitamente es de Clasificación del estado de Ánimo del individuo mediante el análisis de la expresión facial de la persona. Por ende llevaremos adelante la creación de distintas Redes Neuronales, con arquitecturas distintas, para poder evaluar su performance y determinar cuál de ellas se ajusta mejor a nuestra necesidad.

Las arquitecturas de las Redes que llevaremos adelante son:

- I. RNA con 1 Capa Oculta
- II. RNA con 2 Capas Ocultas
- III. RNA con 3 Capas Ocultas
- IV. RNA con 4 Capas Ocultas

Adicionalmente, como valor adicional, incorporaremos una Red Neuronal de Convolución (CNN ya que son redes cuyo uso principal es el del análisis de imágenes. Por ese motivo intentaremos implementarla para compararla con las tradicionales MLP.

Para llevar adelante el modelado, emplearemos una distribución de 75% de datos para Train y el 25% de los datos restantes para Test. De esta manera contamos con la siguiente distribución de datos para emplear las Redes



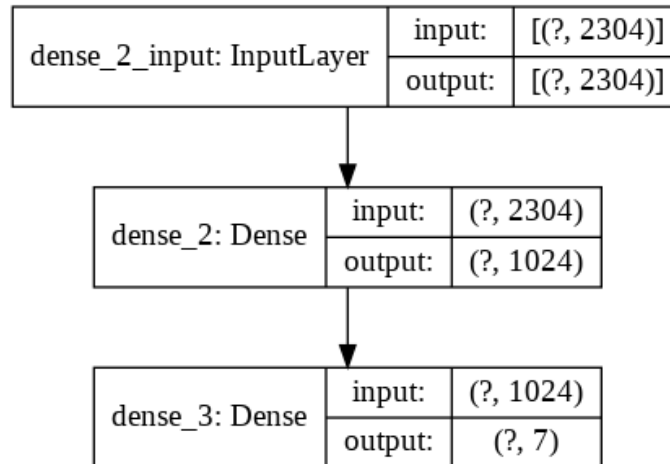
Para unificar criterios en nuestras redes neuronales, setaremos algunas características para todas iguales.

Estas características son:

- Cantidad de Neuronas de Entrada: 2304
- Cantidad de Neuronas de Salida: 7
- Cantidad de Iteraciones: 300
- Batch Size : 128

Red Neuronal Artificial con 1 Capa Oculta

La arquitectura empleada para esta RNA es la siguiente:



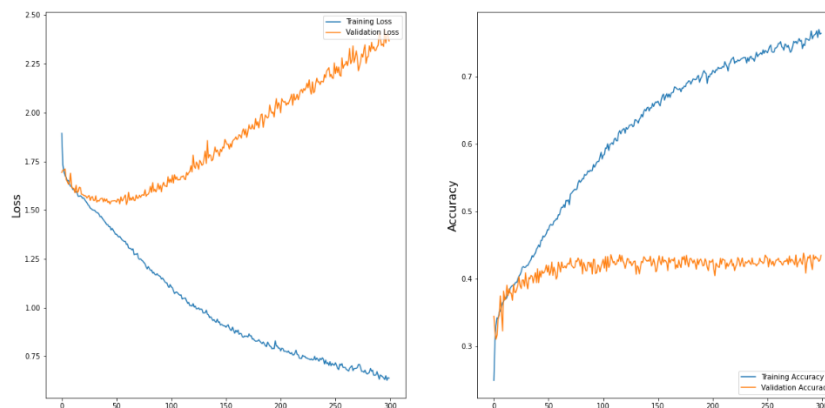
La misma cuenta con:

Capa de Entrada: 2304 neuronas

Capa Oculta: 1024 Neuronas

Capa Salida: 7 Neuronas

Los resultados preliminares de nuestro modelo desarrollado son:



Podemos observar que nuestra red de entrenamiento a medida que van avanzando las iteraciones va aumentando su accuracy y disminuyendo su error (Loss). Mientras que cuando lo constata contra nuestro set de Validación podemos apreciar que el Error a partir de la iteración 50, aproximadamente, el error empieza a incrementarse y el accuracy se comienza a mantener constante (entre 40-43).

Esto nos da el pie para considerar que empieza a ocurrir Overfitting en donde nuestra Red diseñada empieza a aprenderse tan bien nuestro set de Entrenamiento lo que provoca que no pueda aplicarse dicho conocimiento con los datos de validación, y eso se ve reflejado en el Accuracy fluctuante desde la iteración 50 en adelante, combinándolo con el incremento del Error.

A modo de resumen podemos mencionar que nuestro error general de la red es de 2.367 y nuestra Accuracy de 0.435(43,5%). Mientras que los Accuracy y Errores obtenidos en nuestro set de entrenamiento son 0.851(85.1%) y 0.449 respectivamente.

A continuación, evaluaremos que información enriquecedora nos provee las métricas de Precision, Recall, F1-Score, y nuestra Matriz de Confusión.

Emotion	Precision	Recall	F1- Score	Support	Confusion Matrix								
Angry	0.3419	0.2472	0.2869	1238	Actual	Angry	306	10	149	276	160	84	253
Disgust	0.4167	0.2555	0.3167	137		Disgust	17	35	17	30	12	9	17
Fear	0.3333	0.2523	0.2872	1280		Fear	149	8	323	224	165	153	258
Happy	0.5377	0.6564	0.5912	2247		Happy	129	9	104	1475	142	84	304
Sad	0.3811	0.2837	0.3253	1519		Sad	134	10	177	307	431	89	371
Surprise	0.5408	0.5964	0.5672	1001		Surprise	45	3	89	113	47	597	107
Neutral	0.3597	0.4748	0.4093	1550		Neutral	115	9	110	318	174	88	736
Macro AVG	0.4159	0.3952	0.3977	8972									
Weighted AVG	0.4228	0.4350	0.4225	8972									

En base a esta información expuesta, podemos mencion lo siguiente:

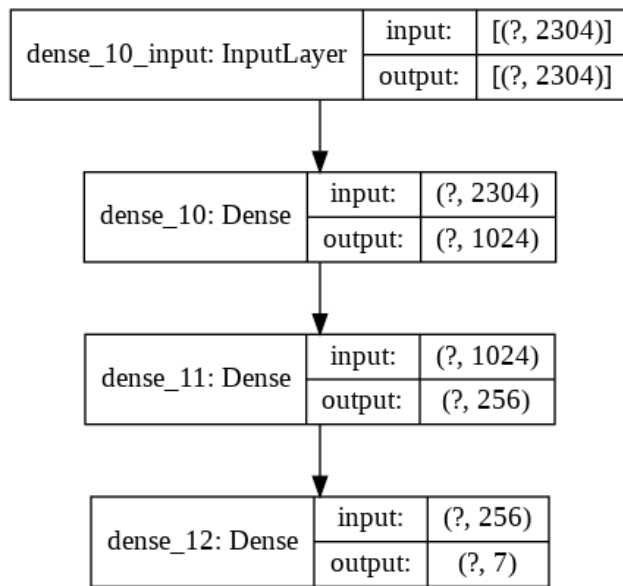
- La emoción que mejor se identifica y clasifica es la de Happy, en donde tiene un Recall (Exhaustividad) del 65,64%. Esto nos indica que de todas las emociones de Felicidad 2247, el 65% es predecida, es decir 1475. Mientras que disponemos de una precisión del 54%, lo que nos indica que del total de la emoción predecida como Happy, solamente ese % es realmente ese estado de ánimo.
- Por otro lado, la emoción Surprise tiene un Recall mayor al 50%, lo cual nos garantiza un acierto mayor a la mitad, con un rango de casi el 60% y con una Precisión superior al 54%.
- El resto de las emociones detectadas no nos garantiza fiabilidad en los datos porque presentan valores inferiores al 50%

En resumen podemos indicar que nuestro modelo sería ineficiente para resolver problemáticas que requieran un análisis preciso de las emociones, si quisiéramos ponderar las emociones de Happy y Surprise, podría ser utilizada esta Red para esta problemática.

Dado que nuestra Exactitud es del 43,5% no podemos decir que nuestro modelo cumpla con las necesidades básicas que el negocio requiera y es necesario emplear nuevas estrategias o mejoras a llevar adelante.

Red Neuronal Artificial con 2 Capas Oculta

La arquitectura empleada para esta RNA es la siguiente:



La misma cuenta con:

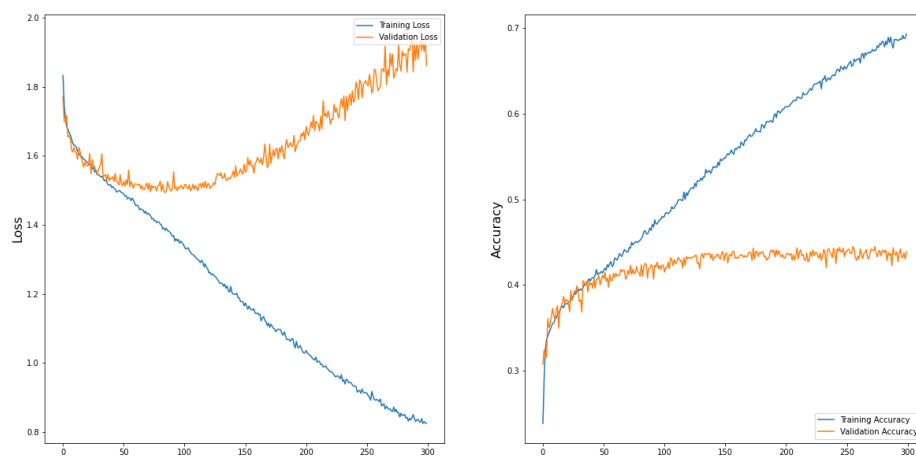
Capa de Entrada: 2304 neuronas

Primer Capa Oculta: 1024 Neuronas

Segunda Capa Oculta: 256 Neuronas

Capa Salida: 7 Neuronas

Los resultados básicos de nuestra Red Neuronal de 2 Capas Ocultas son los siguientes:



Podemos observar que nuestra red con los datos de entrenamiento a medida que van avanzando las iteraciones va aumentando su accuracy y disminuyendo su error (Loss). Mientras que cuando lo constata contra nuestro set de Validación podemos apreciar que el Error a partir de la iteración 125, aproximadamente, el error empieza a incrementarse y el accuracy se comienza a mantener constante (entre 40-43). Esto nos da el pie para considerar que empieza a ocurrir Overfitting en donde nuestra Red diseñada empieza a aprenderse tan bien nuestro set de Entrenamiento lo que provoca que no pueda aplicarse dicho conocimiento con los datos de validación, y eso

se ve reflejado en el Accuracy fluctuante desde la iteración 125 en adelante, combinándolo con el incremento del Error.

A diferencia de la red diseñada anteriormente, este overfitting empieza a ocurrir en iteraciones posteriores, producto de la arquitectura diseñada (2 Capas Ocultas vs 1 Capa Oculta) dado que al tener mas neuronas intermedias y capas, se requieren mas Epochs para que la red aprenda de memoria los datos.

A modo de resumen podemos mencionar que nuestro error general de la red es de 1.861 y nuestra Accuracy de 0.439(43,9%).

Mientras que los Accuracy y Errores obtenidos en nuestro set de entrenamiento son 0.799(80%) y 0.622 respectivamente.

A continuación, evaluaremos que información enriquecedora nos provee las métricas de Precision, Recall, F1-Score, y nuestra Matriz de Confusión.

Emotion	Precision	Recall	F1- Score	Support
Angry	0.3348	0.3069	0.3203	1238
Disgust	0.4828	0.3066	0.3750	137
Fear	0.3324	0.2852	0.3070	1280
Happy	0.5811	0.6093	0.5948	2247
Sad	0.3391	0.4128	.3723	1519
Surprise	0.5420	0.6064	0.5724	1001
Neutral	0.4130	0.3535	0.3810	1550
Macro AVG	0.4322	0.4115	0.4175	8972
Weighted AVG	0.4357	0.4389	0.4354	8972

		Confusion Matrix						
Actual	Angry	380	10	148	196	263	78	163
	Disgust	22	42	16	20	20	7	10
	Fear	156	7	365	160	278	167	147
	Happy	174	4	161	1369	273	89	177
	Sad	188	9	156	236	627	75	228
	Surprise	68	3	105	95	69	607	54
	Neutral	147	12	147	280	319	97	548
		Angry	Disgust	Fear	Happy Predicted	Sad	Surprise	Neutral

En base a esta información expuesta, podemos mencion lo siguiente:

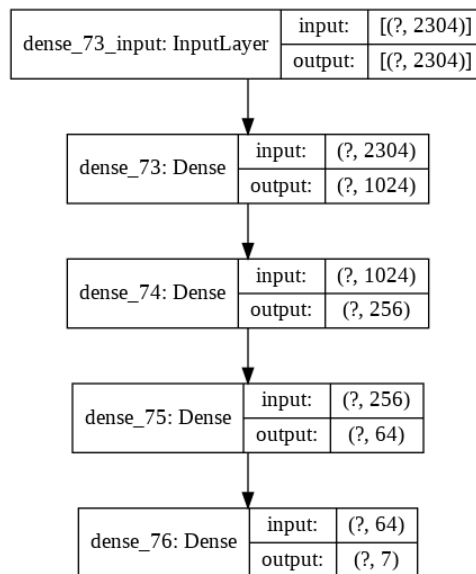
- La emociones de Happy y Surprise tienen rendimientos similares, con un Recall similar y una Precisión mayor a 50%.
- Para el resto de las emociones vemos que aun no conseguimos valores representativos para considerar la red útil, dado que no superan el 50% de Precision ni de Exhaustividad

Como conclusión, podemos mencionar que el rendimiento entre esta red y la anterior, a simple vista parecen idénticos, pero debemos decir que esta red es sutilmente superior a la otra.

Presenta un Accuracy mayor y un error menor. Así mismo la precisión y Exhaustividad también es superior, por lo que si tuviésemos que elegir entre esta red y la anterior, deberíamos quedarnos con esta.

Red Neuronal Artificial con 3 Capas Oculta

La arquitectura empleada para esta RNA es la siguiente:



La misma cuenta con:

Capa de Entrada: 2304 neuronas

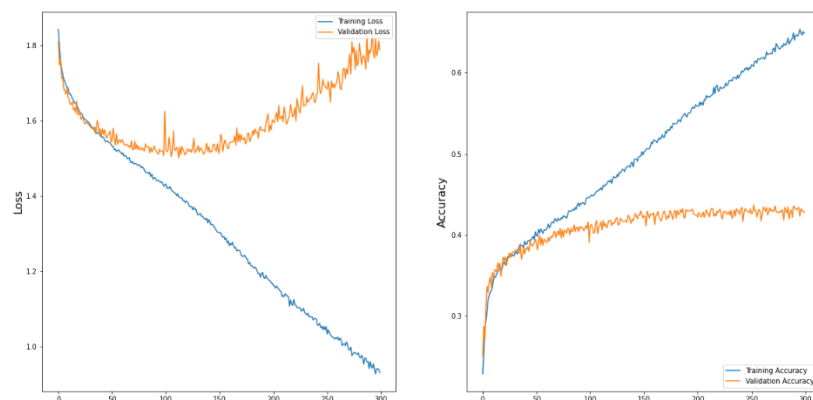
Primer Capa Oculta: 1024 Neuronas

Segunda Capa Oculta: 256 Neuronas

Tercer Capa Oculta: 64 Neuronas

Capa Salida: 7 Neuronas

Los resultados básicos de nuestra Red Neuronal de 3 Capas Ocultas son los siguientes:



Como ya venimos mencionando anteriormente, se observa que a partir de la Iteración 150, el error en nuestro set de validación comienza a incrementarse y nuestro accuracy aplanarse. Esto nos da la sensación de que nuestra empieza a sobreajustar sobre el set de entrenamiento y el accuracy no incrementaría mucho más a partir de ese momento. Si bien vemos que en nuestro set de entrenamiento incrementa, se debe a que va aprendiendo los datos de memoria y esto tiene una correlación con el error para el mismo conjunto de datos.

A modo de resumen podemos mencionar que nuestro error general de la red es de 1.787 y nuestra Accuracy de 0.428(42,8%).

Mientras que los Accuracy y Errores obtenidos en nuestro set de entrenamiento son 0.741 (74,1%) y 0.734 respectivamente.

A continuación, evaluaremos que información enriquecedora nos provee las métricas de Precision, Recall, F1-Score, y nuestra Matriz de Confusión:

					Confusion Matrix								
Emotion	Precision	Recall	F1- Score	Support									
Angry	0.3309	0.2609	0.2918	1238	Actual	Angry	323	3	155	211	305	27	214
Disgust	0.7200	0.1314	0.2222	137		Disgust	33	18	17	24	23	3	19
Fear	0.3175	0.2711	0.2925	1280		Fear	141	1	347	168	323	94	206
Happy	0.5868	0.6017	0.5942	2247		Happy	155	2	123	1352	329	44	242
Sad	0.3083	0.4292	0.3588	1519		Sad	157	0	168	219	652	32	291
Surprise	0.6798	0.4985	0.5752	1001		Surprise	37	0	157	88	114	499	106
Neutral	0.3751	0.4174	0.3951	1550		Neutral	130	1	126	242	369	35	647
Macro AVG	0.4741	0.3729	0.3900	8972									
Weighted AVG	0.4418	0.4278	0.4274	8972									
						Angry	Disgust	Fear	Happy Predicted	Sad	Surprise	Neutral	

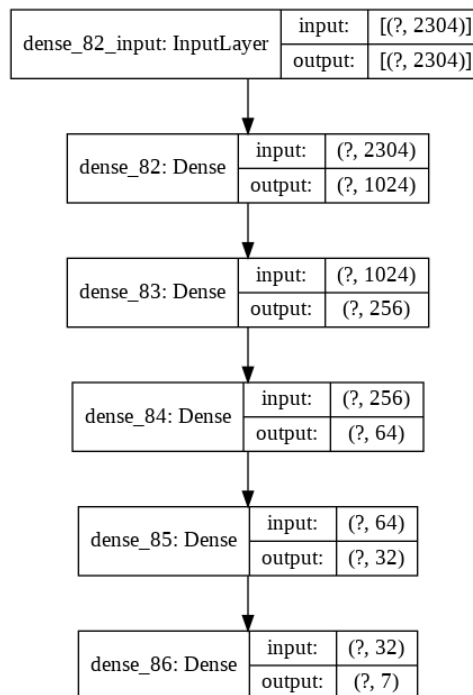
En base a esta información expuesta, podemos mencion lo siguiente:

- Continuamos detectando que la Emoción que mejor se identifica y clasifica es la de Felicidad con un Recall del 60% y una Precisión del 59%
- Observams que aparece una Clase que antes no figuraba que es la de Disgusto. Esta red logra identificar al 72% de esta emoción, pero como punto negativo tiene una Exhaustividad demasiado baja con un 13%. Lo que nos dice es que realmente no acierta con la predicción constatándolo con la emoción real.

En resumen podemos identificar que está red es inferior a las dos anteriormente construidas en lo que representa Recall y Accuracy. Pero logramos detectar que esta Red tiene una Precisión mayor a las otras. Esto quiere decir que logra detectar un mayor caso de emociones al momento de predecirlas, pero al final realmente no logra ser asertiva con esas predicciones porque tiene una Exhaustividad inferior.

Red Neuronal Artificial con 4 Capas Oculta

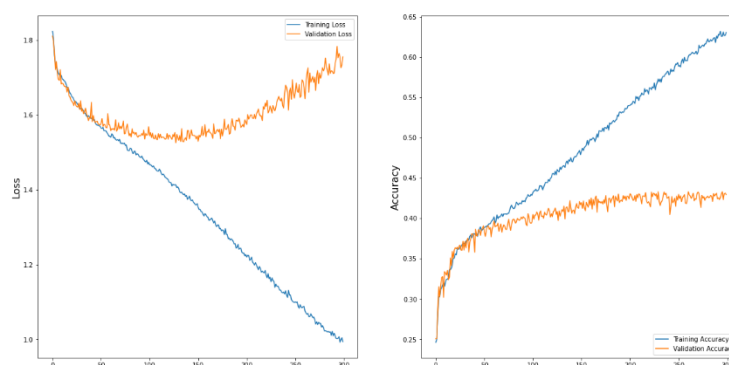
La arquitectura empleada para esta RNA es la siguiente:



La misma cuenta con:

- Capa de Entrada: 2304 neuronas
- Primer Capa Oculta: 1024 Neuronas
- Segunda Capa Oculta: 256 Neuronas
- Tercer Capa Oculta: 64 Neuronas
- Cuarta Capa Oculta: 32 Neuronas
- Capa Salida: 7 Neuronas

Los resultados básicos de nuestra Red Neuronal de 4 Capas Ocultas son los siguientes:



Como ya venimos mencionando anteriormente, se observa que a partir de la Iteración 200, el error en nuestro set de validación comienza a incrementarse y nuestro accuracy aplanarse. Esto nos da la sensación de que nuestra empieza a sobreajustar sobre el set de entrenamiento y el accuracy no incrementaría mucho más a partir de ese momento. Si bien vemos que en nuestro set de entrenamiento incrementa, se debe a que va aprendiendo los datos de memoria y esto tiene una correlación con el error para el mismo conjunto de datos.

A diferencia de las Redes anteriores, al tener mayor Capas Ocultas, va necesitando mas iteraciones para comenzar a sobreajustar y aprender de memoria los datos.

A modo de resumen podemos mencionar que nuestro error general de la red es de 1.755 y nuestra Accuracy de 0.429(42,9%).

Mientras que los Accuracy y Errores obtenidos en nuestro set de entrenamiento son 0.702 (70,2%) y 0.816 respectivamente.

A continuación, evaluaremos que información enriquecedora nos provee las métricas de Precision, Recall, F1-Score, y nuestra Matriz de Confusión:

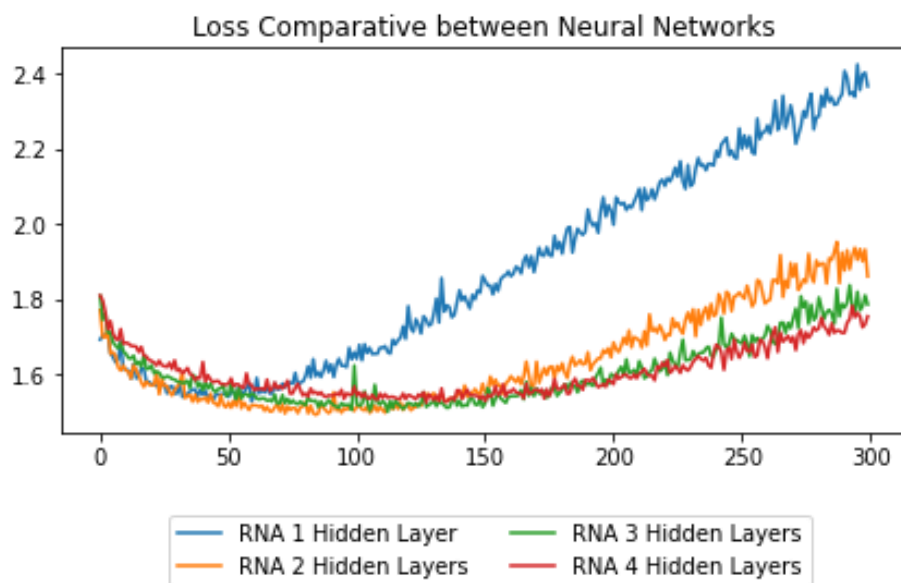
Emotion	Precision	Recall	F1- Score	Support	Confusion Matrix								
Angry	0.3078	0.2851	0.2960	1238	Actual \ Predicted	Angry	353	8	137	341	144	29	226
Disgust	0.4571	0.1168	0.1860	137		Disgust	37	16	15	43	7	3	16
Fear	0.3114	0.2148	0.2543	1280	Fear	195	5	275	342	141	102	220	
Happy	0.4814	0.7481	0.5858	2247		Actual \ Happy	151	2	91	1681	97	29	196
Sad	0.4032	0.2482	0.3073	1519	Sad	194	0	153	434	377	29	332	
Surprise	0.6721	0.4935	0.5691	1001		Surprise	57	0	116	201	34	494	99
Neutral	0.3759	0.4235	0.3982	1550	Neutral	160	4	96	450	135	49	656	
Macro AVG	0.4299	0.3614	0.3710	8972		Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral	
Weighted AVG	0.4226	0.4293	0.4110	8972									

En base a esta información expuesta, podemos mencion lo siguiente:

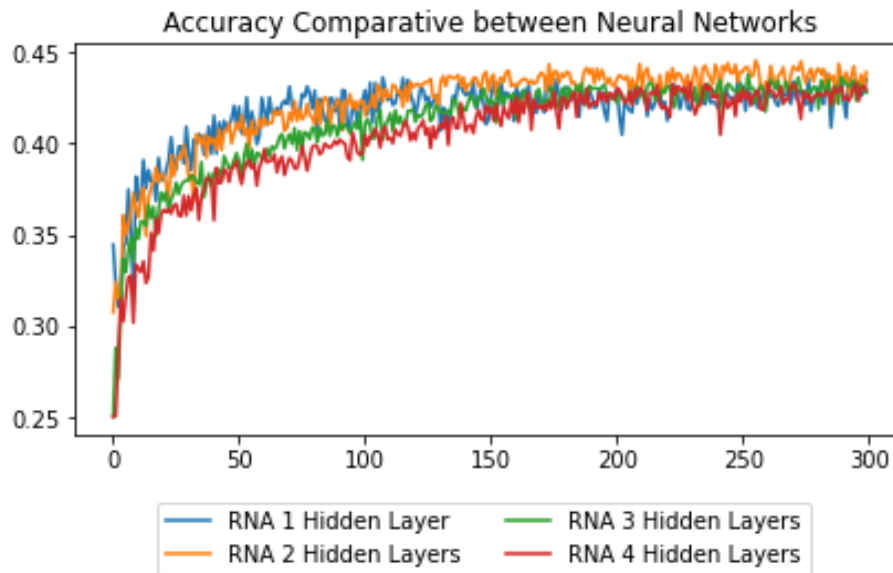
- Detectamos que esta Red trabaja muy bien con la emoción Happy, ya que tiene una performance de casi el 75%, eso quiere decir que de los que están Felices predice el 75% de ellos correctamente. No obstante solo logra identificar al 48%.
- Esta red logra identificar en un 67% a los individuos que están Sorprendidos, pero sobre ellos tiene un acierto o performance del 50%

3. Resultados

Para el análisis de resultados, realizaremos una evaluación comparativa entre nuestros modelos para determinar cuál de ellos cubre nuestras necesidades, en función a las métricas conocidas.



Se aprecia brevemente que la Red de 1 sola Capa Oculta es la que presenta a lo largo de las iteraciones un mayor incremento de error en el Modelo. Esto se puede dar por el hecho de que al solo tener 1 sola capa extra además de la de Entrada y Salida, las chaces de lograr un mayor acierto es más complejo.



Para el análisis de la Exactitud, se observa que el comportamiento de las 4 Redes a lo largo del tiempo se mantiene similar. Todas representan un crecimiento similar, pero en todas se aprecia que superada la iteración 200 ya el crecimiento es lineal y no se nota un crecimiento exponencial. Ninguna de estas redes tiene un Accuracy superior al 45%

A modo resumen podemos realizar el siguiente cuadro comparativo:

RNA MLP	Accuracy	Val Accuracy	Loss	Val Loss
1 Capa Oculta	85,10%	43,50%	0.449	2.367
2 Capas Ocultas	79,90%	43,90%	0.622	1.861
3 Capas Ocultas	74,10%	42,80%	0.734	1.787
4 Capas Ocultas	70,20%	42,90%	0.816	1.755

En base a este cuadro podríamos determinar, en caso de tener que elegir un modelo de los 4 contruidos, el mejor sería el segundo, con 2 capas ocultas. Presenta una mejor Accuracy sobre nuestros datos de validación, y con un error de 1.861.

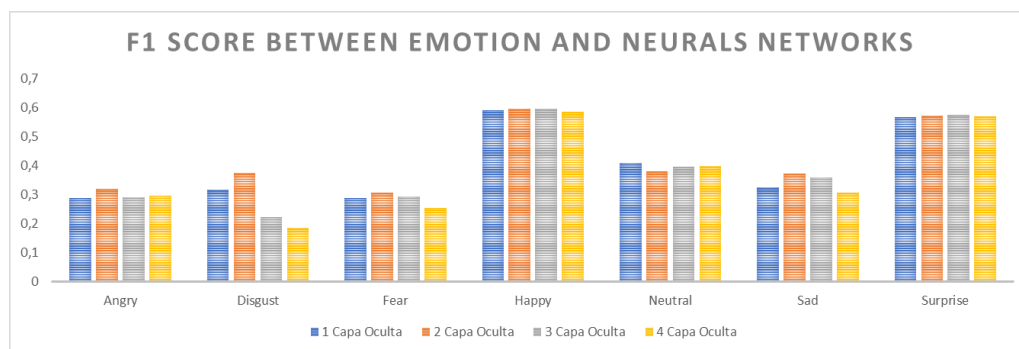
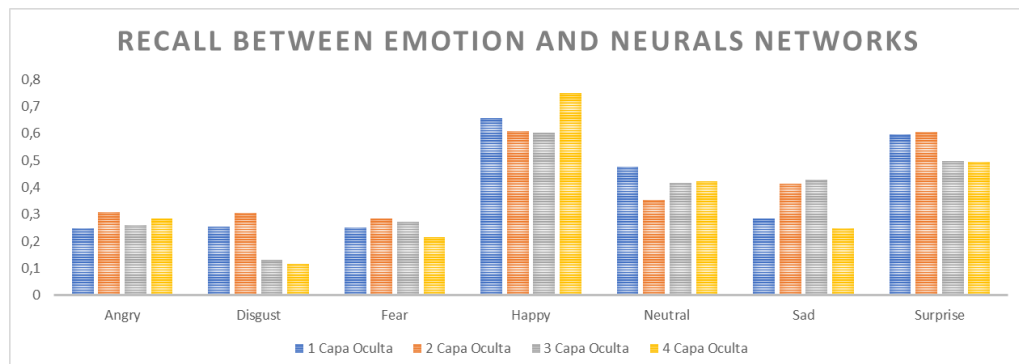
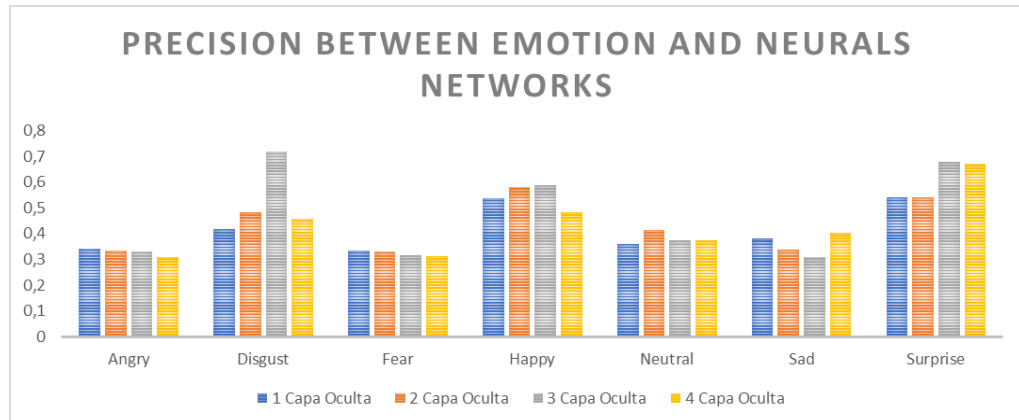
Adicionalmente vemos que a medida que agregamos Capas Ocultas nuestra Accuracy de Entrenamiento va disminuyendo y nuestro Error Aumentando, esto es porque le lleva más tiempo efectuar el sobreajuste sobre los datos por su arquitectura.

Si complementamos el análisis con los datos brindados por nuestro reporte de clasificación vemos que:

RNA MLP	Precision	Recall	F1 Score
1 Capa Oculta	42,28%	43,50%	42,25%
2 Capas Ocultas	43,57%	43,89%	43,54%
3 Capas Ocultas	44,18%	42,78%	42,74%
4 Capas Ocultas	42,26%	42,93%	41,10%

Se sigue validando que la mejor red de las 4, sigue siendo aquella con 2 Capas Ocultas, dado que presenta un mayor valor de F1 Score, que representa una media entre los otros dos indicadores. No obstante si quisiéramos considerar una red con mayor precisión la elección sería la Red de 3 Capas Ocultas.

Mas allá de estos valores, como la diferencia entre las redes neuronales es ínfima, podemos también tomar en consideración estos mismos indicadores pero desde el punto de vista de cada una de las clases, para ver cual realmente cubre nuestras necesidades. Dado que si nuestro objetivo es priorizar que se clasifique mejor la emoción de Felicidad por sobre el resto, deberíamos evaluar que red presenta un mayor valor de Recall y Precisión para esta. Por este motivo procedemos a evaluar cada clase por separado comparando cada una de las redes.



De esta manera tenemos una mirada Cross emociones y poder evaluar cuál de las redes ponderaríamos.

Por ejemplo si quisiéramos que nuestra red sea más eficiente con la emoción de Happy, tendríamos que tomar la red de 4 Capas Ocultas ya que nos brinda un 75% de Exhaustividad, porque nos garantiza que del 100% de personas que están felices, esta red es capaz de detectar el 75% de ellas.

Por otro lado si nuestro nicho es de detectar gente que no esté feliz, sino triste o emociones similares, la evaluación o análisis sería similar. Para este ejemplo se tomaría la Red de 3 Capa Oculta porque tiene mayor porcentaje de Recall en la emoción Sad que el resto.

4. Discusión

En base a los resultados expuestos en secciones anteriores, podemos determinar que las redes confeccionadas no resuelven fehacientemente la problemática, ya que disponen de una performance baja. Lo que implica que no se puede tomar la información que nos arrojan estos modelos como únicos, y es necesario una validación o revisión desde el punto de vista del experto.

Los factores que podemos entender que provoca este inconveniente pueden ser varias:

- La calidad de las imágenes es baja para entrenar esta red.
- El dataset utilizado para entrenar y validar el modelo no es bueno
- La arquitectura utilizada para resolver este problema, puede que no sea la mejor. Se podrían utilizar otros tipos de redes neuronales para incrementar la performance
- Las personas muchas veces no expresan sus emociones en bases a gestos, sino que asimilan sus problemas de manera internamente, por lo que muchas veces captar el estado de ánimo de la persona en base a una foto puede que sea lo suficiente complejo.

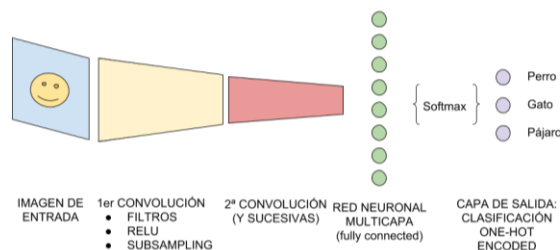
En base a esto creemos que es necesario abordar la problemática mediante el uso de otra arquitectura disponible.

A continuación diseñaremos una arquitectura de redes diferente llamada Red Neuronal Convolutiva (CNN) para ver si mejora nuestra performance, con respecto a nuestras redes Back Propagation, y nos brinde una alternativa superior para la toma de decisiones.

5. Redes Neuronales Convolucionales (CNN)

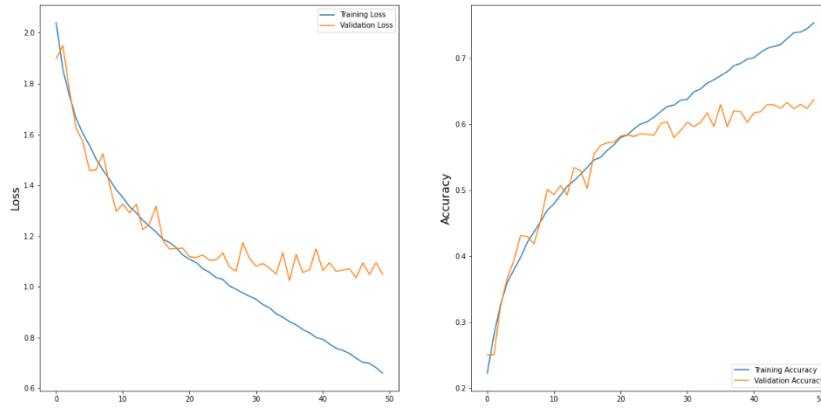
La CNN es un tipo de Red Neuronal Artificial con aprendizaje supervisado que procesa sus capas imitando al cortex visual del ojo humano para identificar distintas características en las entradas que en definitiva hacen que pueda identificar objetos y “ver”. Para ello, la CNN contiene varias capas ocultas especializadas y con una jerarquía: esto quiere decir que las primeras capas pueden detectar líneas, curvas y se van especializando hasta llegar a capas más profundas que reconocen formas complejas como un rostro o la silueta de un animal.

La arquitectura de una CNN es similar a:



Dado que nuestro problema se basa básicamente en detectar la imagen y poder clasificarla, creemos que esta Red se ajusta mucho más a nuestra problemática. A continuación veremos los resultados obtenidos de esta red aplicando los datasets que disponemos.

Para esta Red empleamos 50 iteraciones (Epochs).



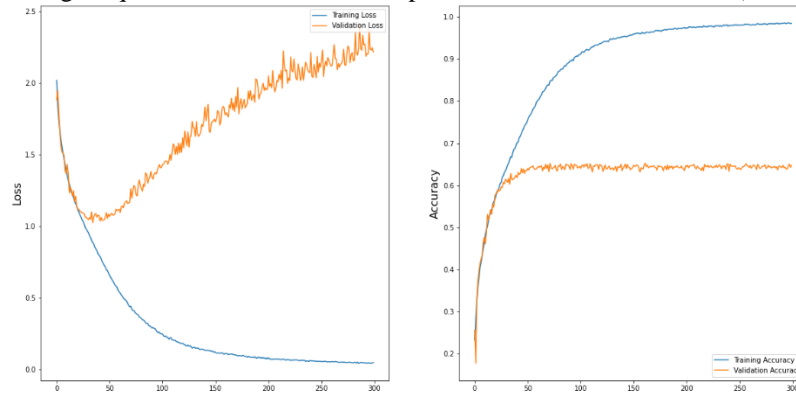
En un primer análisis vemos que nuestro accuracy se va normalizando a partir del 60%.

Nuestra Red nos presenta los siguientes datos:

Accuracy: 0.638

Loss: 1.049

Si entrenamos nuestra red con 300 iteraciones vemos que nuestro Accuracy(0,647) se mantiene igual que con menos iteraciones, pero nuestro Error incrementa (2.215).



Esto ocurre porque nuestra Red Neuronal aprende demasiado bien nuestros datos de entrenamiento, lo cual genera que al momento de validar la Red contra los datos de Prueba, el error es mayor por el sobreajuste. Por ende no es necesario siempre tener un Accuracy tan alto en nuestro dataset de prueba, si nuestro accuracy de validación se mantiene igual.

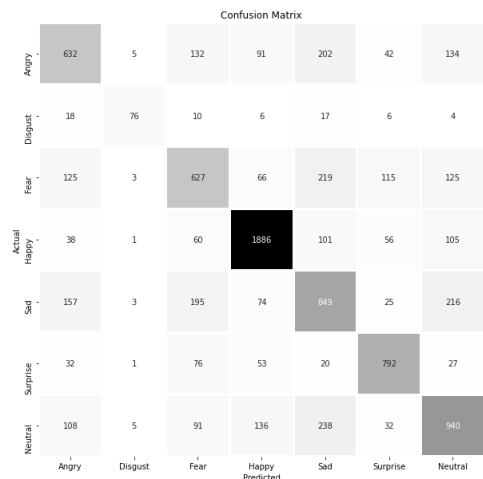
En este ejemplo vemos que nuestra Performance es la misma para 50 que para 300 iteraciones, pero con un error menor en la de Epochs menores.

Con esto podemos definir que nuestra red es bastante superior a las redes tradicionales que antes analizamos, ya que supera su accuracy casi en 20% (44% era la mejor).

En base a esto podríamos determinar que el uso de una CNN sería la mejor opción para este tipo de problemáticas.

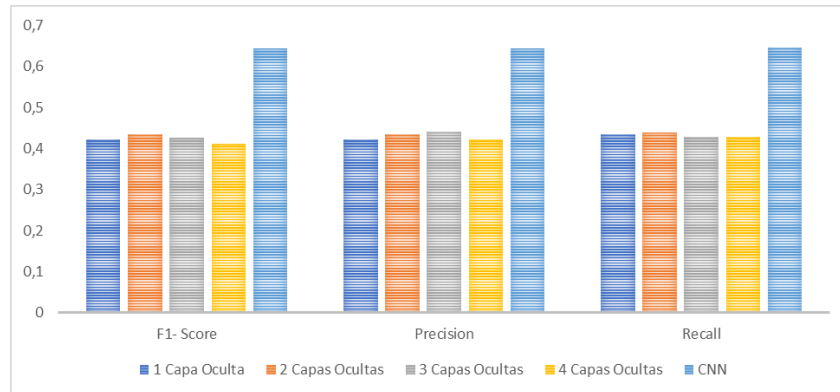
De todas formas evaluaremos que información enriquecedora nos provee las métricas de Precision, Recall, F1-Score, y nuestra Matriz de Confusión:

Emotion	Precision	Recall	F1- Score	Support
Angry	0.5694	0.5105	0.5383	1238
Disgust	0.8085	0.5547	0.6580	137
Fear	0.5264	0.4898	0.5075	1280
Happy	0.8157	0.8393	0.8274	2247
Sad	0.5158	0.5589	0.5365	1519
Surprise	0.7416	0.7912	0.7656	1001
Neutral	0.6061	0.6065	0.6063	1550
Macro AVG	0.6548	0.6216	0.6342	8972
Weighted AVG	0.6451	0.6467	0.6449	8972

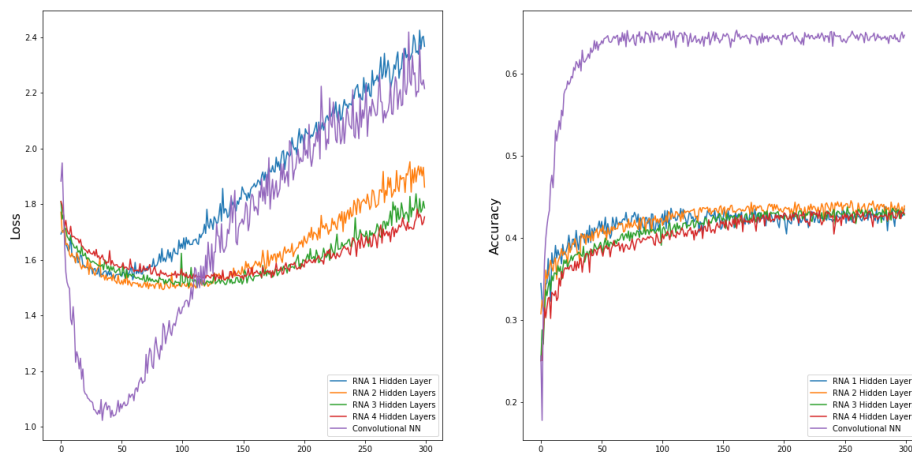


A simple vista podemos determinar que casi todas las clases tienen un % de acierto superior al 50%, lo que nos brinda una mayor oportunidad de acierto al momento de clasificarlas.

Adicionalmente si vemos las variables de Precision, Recall y F1 Score nos va a arrojar valores altamente superiores que las otras redes, como vemos en el gráfico siguiente:



Adicionalmente podemos decir que la emoción con inconvenientes de clasificar es la clase Fear, que puede confundirse con la clase Sad. Esto puede darse por la calidad de la imagen, una similitud entre los rasgos con la Clase mencionada, o porque el dataset no es bueno.








Con estos gráficos comparativos, vemos como la Red CNN es superior en el accuracy, y el error que nos brinda depende mucho de cuantas Iteraciones decidamos utilizar para la red.

Vemos que si consideramos un tope de 50 iteraciones, el error es el mínimo y el Accuracy es constante a partir de ese valor. Por lo que esta red con Epochs de 50 es la decisión más acertada.

6. Casos Prácticos

A continuación veremos como se comportan nuestras redes con situaciones de la cotidianidad, para entender si pueden ser utilizadas para la toma de decisiones.

Imagen a analizar	Topología de Red	Emoción Predecida
	MLP 1 capa Oculta	Happy
	MLP 2 Capas Ocultas	Angry
	MLP 3 Capas Ocultas	Happy
	MLP 4 Capas Ocultas	Happy
	CNN	Happy
	MLP 1 capa Oculta	Sad
	MLP 2 Capas Ocultas	Neutral
	MLP 3 Capas Ocultas	Happy
	MLP 4 Capas Ocultas	Happy
	CNN	Sad
	MLP 1 capa Oculta	Fear
	MLP 2 Capas Ocultas	Angry
	MLP 3 Capas Ocultas	Sad
	MLP 4 Capas Ocultas	Neutral
	CNN	Neutral
	MLP 1 capa Oculta	Angry
	MLP 2 Capas Ocultas	Sad
	MLP 3 Capas Ocultas	Surprise
	MLP 4 Capas Ocultas	Fear
	CNN	Surprise
	MLP 1 capa Oculta	Neutral
	MLP 2 Capas Ocultas	Angry
	MLP 3 Capas Ocultas	Angry
	MLP 4 Capas Ocultas	Happy
	CNN	Neutral

Como podemos ver, los resultados varían entre las redes. Esto ocurre por el entrenamiento a los que cada una fue expuesta. Si todas tuviesen un grado alto de Accuracy, probablemente nos arrojarían similares valores predichos entre las redes. Esto nos deja en evidencia que para poder implementar una Red Neuronal de esta índole, es necesario un trabajo más exhaustivo sobre la información disponible para garantizar una buena performance y obtener una mayor cantidad de acierto sobre la emoción de una persona mediante el reconocimiento de expresiones faciales.

7. Conclusiones

Las redes neuronales son una herramienta muy potente y útil para problemáticas que impliquen toma de decisiones, predicciones o situación de clasificación. De todas formas su uso depende mucho de la situación que se busca resolver y la información que desea obtener el usuario final. Si el usuario además de la predicción busca también información que le provea el modelo, en ese caso se debería buscar otras opciones vigentes que hay en el mercado (Regresión Lineal, Logística, Árbol de decisión, etc).

Por estos motivos, creemos que si bien la mayoría de los problemas se pueden resolver implementando una red neuronal, no siempre es la mejor opción a los problemas, sino que muchas veces debemos evaluar si el modelo debe brindarnos información o solo nos quedamos con el resultado final.

Para nuestro caso en particular, el uso de una Red neuronal resuelve nuestra problemática de manera correcta, ya que nos brinda todo lo necesario para llevar a cabo nuestro objetivo. Pero por todo lo anteriormente mencionado, nuestras redes no fueron lo suficientemente efectivas para resolver la cuestión. En ese caso, si estuviésemos en un ámbito laboral, deberíamos solicitar más información o de mejor calidad para aumentar la performance de nuestras redes neuronales diseñadas.

7. Referencias

1. <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/>
2. <https://www.kaggle.com/yogie25/fer2013>
3. <https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>
4. <https://www.iartificial.net/precision-recall-f1-accuracy-en-clasificacion/>
5. <https://www.digitalocean.com/community/tutorials/how-to-detect-and-extract-faces-from-an-image-with-opencv-and-python>