

Taller 3: Redes Neuronales y SVM.

Docente: Nicolás Abuhadba.

Objetivo

- Predecir si una *startup* que actualmente está en funcionamiento, se convierte en un éxito o en un fracaso.
- El éxito de una empresa se define como el evento que otorga a los fundadores de la empresa una gran suma de dinero a través del proceso de M&A (*Fusión y Adquisición*) o una OPI (*Oferta Pública Inicial*). Una empresa se consideraría fallida si tuviera que cerrar.

Tipo de actividad	Taller grupal
Taller	Redes Neuronales y SVM
Evaluación	Formativa
Tiempo estimado de entrega	20/11/2021

Actividades para desarrollar:

1. Breve contexto

Una *startup* o *start-up* es una empresa o proyecto iniciado por un emprendedor para buscar, desarrollar y validar un modelo económico escalable. Mientras que el espíritu empresarial se refiere a todos los negocios nuevos, incluidos el trabajo por cuenta propia y los negocios que nunca tienen la intención de registrarse, las nuevas empresas se refieren a los nuevos negocios que tienen la intención de crecer más allá del fundador en solitario.

Las empresas emergentes se enfrentan a una gran incertidumbre y tienen altas tasas de fracaso, pero una minoría de ellas llega a ser exitosa e influyente. Algunas *startups* se convierten en unicornios: empresas de nueva creación de propiedad privada valoradas en más de mil millones de dólares.

Las empresas emergentes juegan un papel importante en el crecimiento económico. Aportan nuevas ideas, estimulan la innovación, crean empleo y, por lo tanto, mueven la economía. Ha habido un crecimiento exponencial de nuevas empresas en los últimos años. Predecir el éxito de una *startup* permite a los inversores encontrar empresas con potencial de crecimiento rápido, lo que les permite estar un paso por delante de la competencia.

2. Descripción de datos

Los datos contienen tendencias de la industria, conocimientos de inversión e información de empresas individuales. Hay 48 columnas / características. Algunas de las características son:

- agefirstfunding_year – quantitative
- agelastfunding_year – quantitative
- relationships – quantitative
- funding_rounds – quantitative
- fundingtotalusd – quantitative
- milestones – quantitative
- agefirstmilestone_year – quantitative
- agelastmilestone_year – quantitative
- state – categorical
- industry_type – categorical
- has_VC – categorical
- has_angel – categorical
- has_roundA – categorical
- has_roundB – categorical
- has_roundC – categorical
- has_roundD – categorical
- avg_participants – quantitative
- is_top500 – categorical
- status(acquired/closed) - categórico (la variable objetivo, si una startup es "adquirida" por alguna otra organización, significa que la startup tuvo éxito)

Inspiración

Predecir el éxito de una *startup* permite a los inversores encontrar empresas con potencial de crecimiento rápido, lo que les permite estar un paso por delante de la competencia.

Entregable:

- ✓ Usando la plataforma *Colab* o *RStudio*, elabore un reporte con el desarrollo y creación de un *modelo predictivo* que determine si una *startup* que actualmente está en funcionamiento, se convierte en un éxito o en un fracaso, usando *redes neuronales* y *Maquinas de Soporte Vectorial SVM*.
- ✓ Recuerde que es tan importante el *modelo creado*, como *la explicación de los resultados* paso a paso (no solo código) y una conclusión general.
- ✓ El archivo (*Taller3.pynb* o *Taller3.R*) debe ser subido a la plataforma *Classroom* (pestaña "Evaluación Talleres"). El archivo debe indicar "Taller 3 - Grupo 0x".

Actividades para desarrollar:

- a. Para una red neuronal, busque la *cantidad de neuronas óptimas*, para ello entrene la red neuronal con *una capa oculta* usando 1,2,3,4,5,6,7,8,9,10 neuronas y registre el porcentaje de clasificaciones correctas. Realice 10 veces el entrenamiento para la cantidad de neuronas encontradas, para cada una de las corridas, registre el porcentaje de clasificaciones de la salida de la red. Entregue el porcentaje de clasificaciones correctas al evaluar la red neuronal en el conjunto de prueba. Compare este resultado con el SVM (resultante en ítem b), y discuta cuál tiene mejor rendimiento para predecir si una startup que actualmente está en funcionamiento, se convierte en un éxito o en un fracaso.
- b. Entrene un *support vector machine* usando el conjunto de entrenamiento y entregue el porcentaje de clasificaciones correctas al evaluar el modelo con el conjunto de prueba. En base a la matriz de confusión, que clase le cuesta más predecir al SVM.
- c. Para una red neuronal aleatoria, busque la *cantidad de neuronas óptimas*, y *cantidad de capas ocultas óptimas*; para ello entrene la red neuronal usando 50, 100, 200, 300, 400, 500, 600, 700 neuronas y registre el porcentaje de clasificaciones correctas (*puede usar cualquier función de activación*). Entregue el porcentaje de clasificaciones correctas al evaluar la red neuronal en el conjunto de prueba. Compare este resultado con el SVM, y discuta cuál tiene mejor rendimiento para predecir si una *startup* que actualmente está en funcionamiento, se convierte en un éxito o en un fracaso.