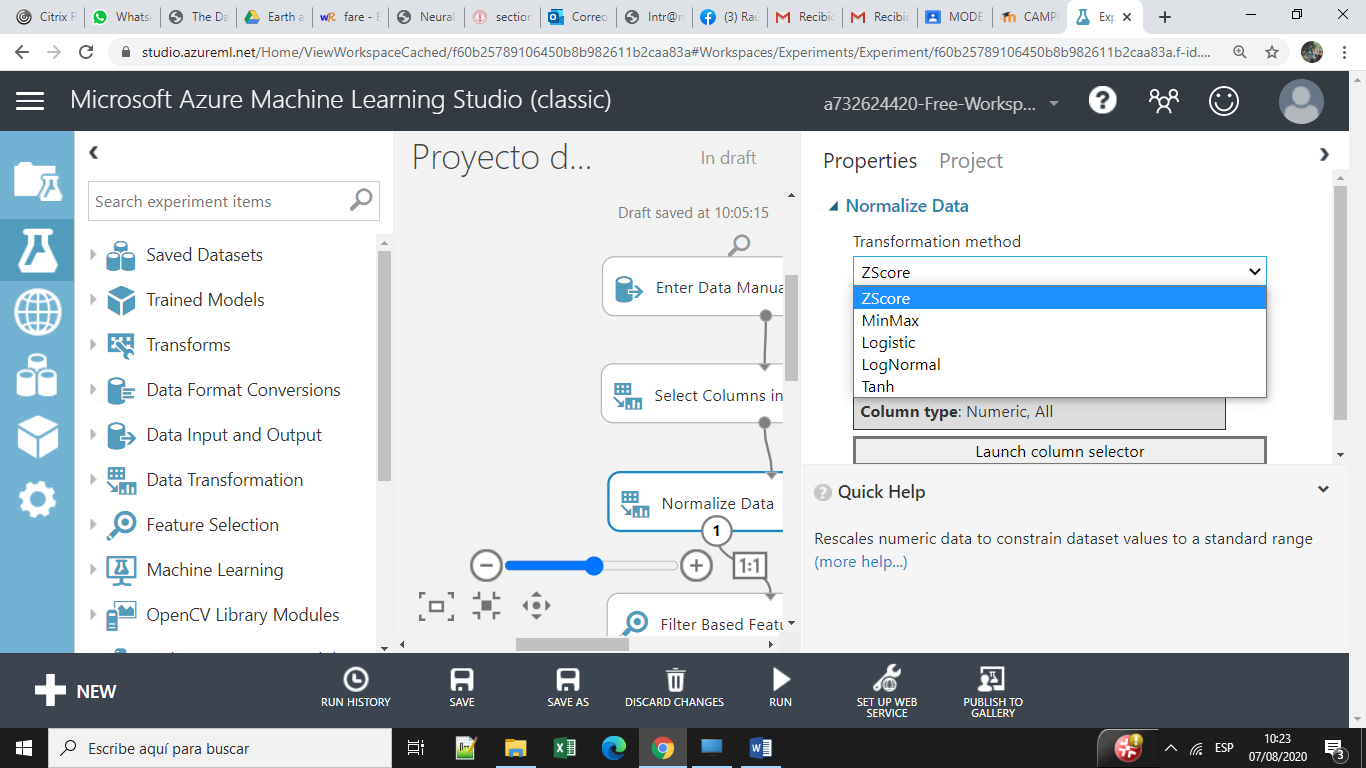
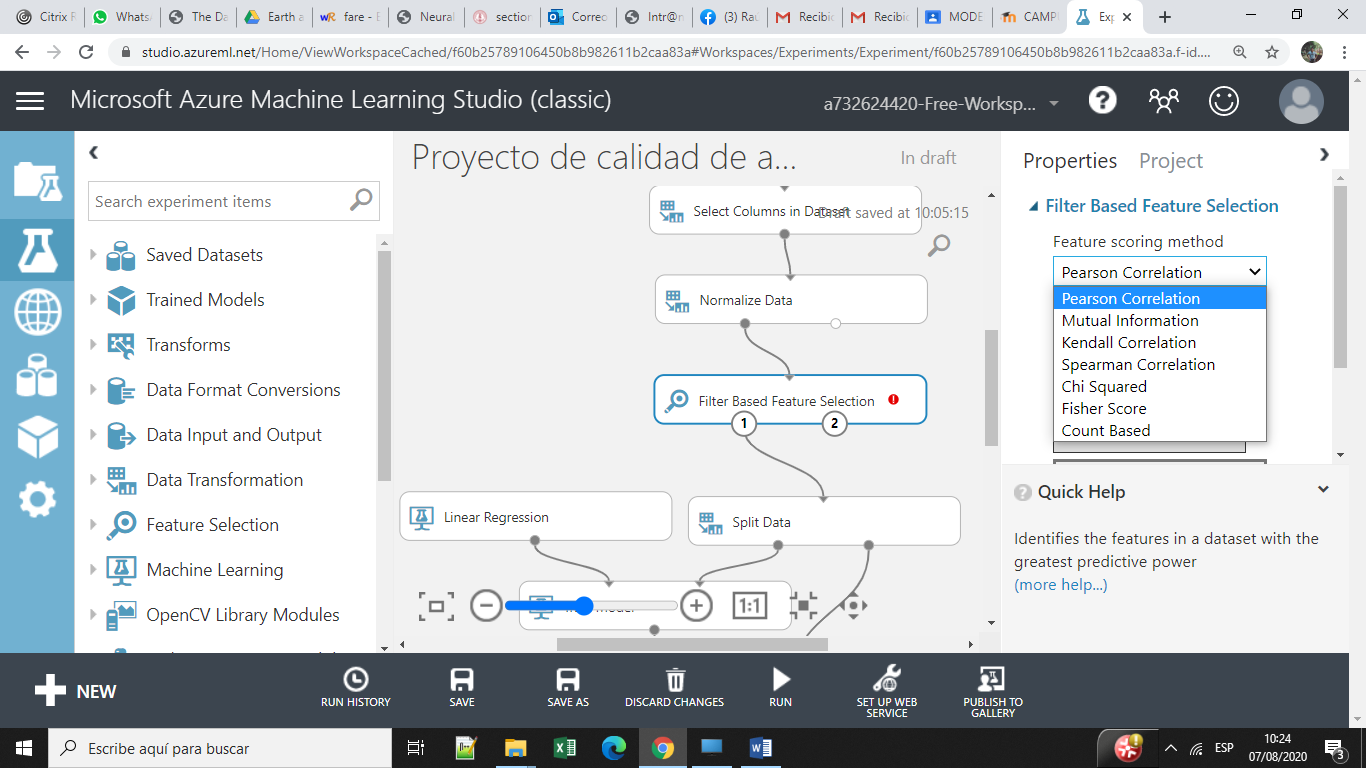
Para normalizar Azure en a la versión libre proporciona los siguiente:





Filter Based Feature Selection

**Feature Selection is one of the core concepts in machine learning which hugely impacts the performance of your model.** The data features that you use to train your machine learning models have a huge influence on the performance you can achieve.

Irrelevant or partially relevant features can negatively impact model performance.

Feature selection and Data cleaning should be the first and most important step of your model designing. Es importante porque:

· **Reduces Overfitting**: Less redundant data means less opportunity to make decisions based on noise.

· **Improves Accuracy**: Less misleading data means modeling accuracy improves.

· **Reduces Training Time**: fewer data points reduce algorithm complexity and algorithms train faster.

* Regression methods are often categorized by the number of response variables. For example, multiple linear regression means a model that has multiple variables to predict.
* In Matlab, multivariate regression refers to a model that has multiple response variables.
* In Azure Machine Learning, regression models support a single response variable.
* In the R language, the features provided for linear regression depend on the package you are using. For example, the **glm** package will give you the ability to create a logistic regression model with multiple independent variables. In general, Azure Machine Learning Studio (classic) provides the same functionality as the R **glm** package.

**CHI-CUADRADO (VARIABLES CUALITATIVAS)**

La prueba de independencia Chi-cuadrado, nos permite determinar si existe una relación entre dos variables categóricas. Es necesario resaltar que esta prueba nos indica si existe o no una relación entre las variables, pero no indica el grado o el tipo de relación; es decir, no indica el porcentaje de influencia de una variable sobre la otra o la variable que causa la influencia.

Por otro lado, Las pruebas chi-cuadrado son un grupo de contrastes de hipótesis que sirven para comprobar afirmaciones acerca de las funciones de probabilidad (o densidad) de una o dos variables aleatorias.

Estas pruebas no pertenecen propiamente a la estadística paramétrica pues no establecen suposiciones restrictivas en cuanto al tipo de variables que admiten, ni en lo que refiere a su distribución de probabilidad ni en los valores y/o el conocimiento de sus parámetros.

Se aplican en dos situaciones básicas:

a)  Cuando queremos comprobar si una variable, cuya descripción parece adecuada, tiene una determinada función de probabilidad. La prueba correspondiente se llama chi-cuadrado de ajuste.

b) Cuando queremos averiguar si dos variables (o dos vías de clasificación) son independientes estadísticamente. En este caso la prueba que aplicaremos ser la chi-cuadrado de independencia o chi-cuadrado de contingencia.

**CORRELACIÓN DE SPEARMAN (VARIABLES CUANTITATIVAS)**

SPEARMAN (Rho de Spearman). Este coeficiente es una medida de asociación lineal que utiliza los rangos, números de orden, de cada grupo de sujetos y compara dichos rangos.

El Coeficiente de Correlación de Spearman, ρ (rho) es una medida de la correlación (la asociación o interdependencia) entre dos variables aleatorias continuas. Para calcular ρ, los datos son ordenados y reemplazados por su respectivo orden. El estadístico ρ viene dado por la expresión: Coeficiente de Correlación de Spearman

Tanto el coeficiente de correlación de Pearson como el de Spearman siguen las mismas normas de interpretación:

- Solamente toma en cuenta valores entre 1 y -1.

- El 0 indica que no existe correlación.

- El valor numérico indica la magnitud de la correlación.

- El coeficiente de correlación cuantifica la correlación entre dos variables, cuando esta realmente existe.

- El hecho de que exista correlación entre las variables no implica que exista causalidad o dependencia entre ellas.

- El signo indica la dirección de la correlación.

- Los valores cercanos a 1 nos indican una correlación muy buena y los cercanos a cero una correlación mínima o nula.

# Q LEARNING

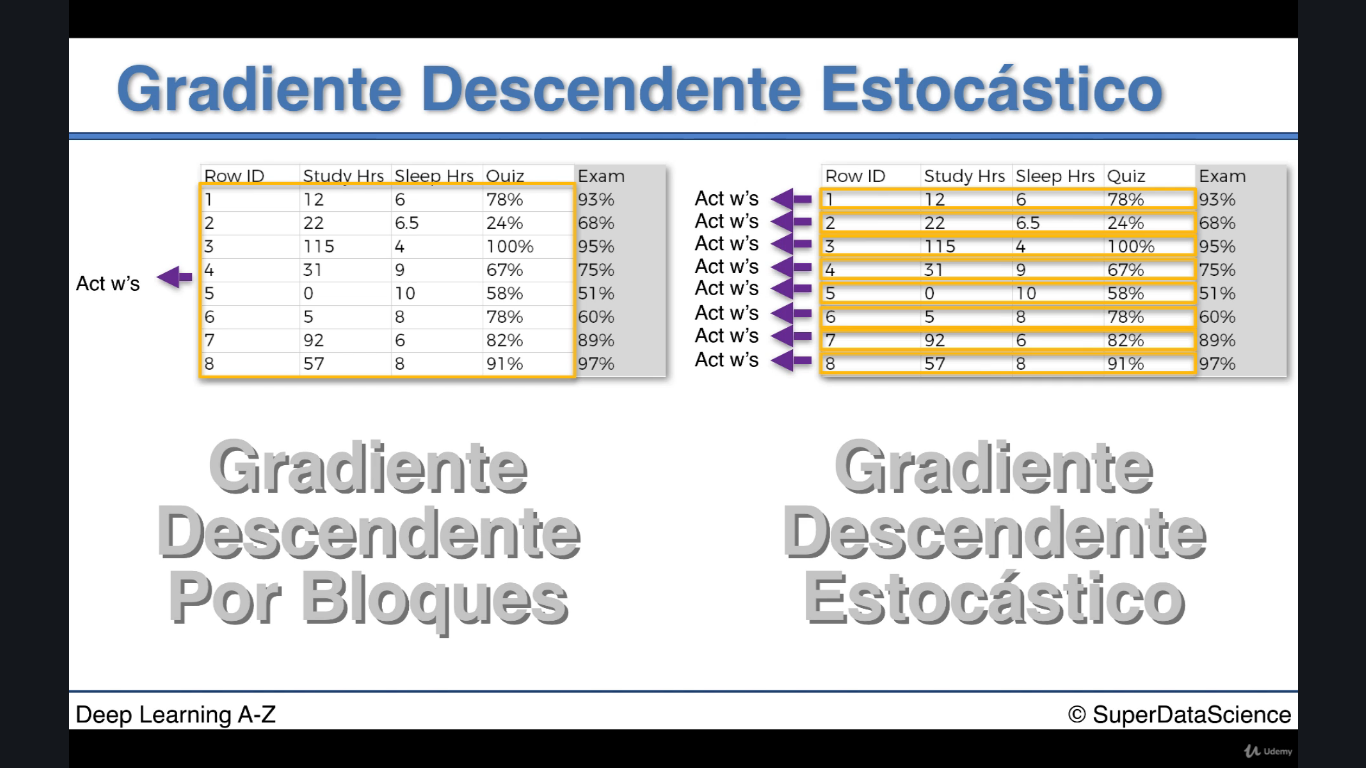
La política

Deep learning and reinforcement learning are both systems that learn autonomously. The difference between them is that deep learning is learning from a training set and then applying that learning to a **new data set**, while reinforcement learning **is dynamically learning by adjusting actions based in continuous feedback to maximize a reward**.

Deep learning and reinforcement learning aren’t mutually exclusive. In fact, you might use deep learning in a reinforcement learning system, which is referred to as deep reinforcement learning and will be a topic I cover in another post. You might also like to explore the [difference between data mining and machine learning](https://bernardmarr.com/default.asp?contentID=1741).

# DEEP Q LEARNING

El método gradiente descendente por bloques realiza una corrección global del experimento. Se puede utilizar ambos simultáneamente.



Políticas de selección de acción

E greedy

E soft ( 1-e)

Softmax