

---

# Módulo 5: Técnicas Avanzadas

# Agenda

- Optimización de hiperparámetros
- Validación Cruzada
- Reducción de dimensionalidad

# Optimización de hiperparámetros



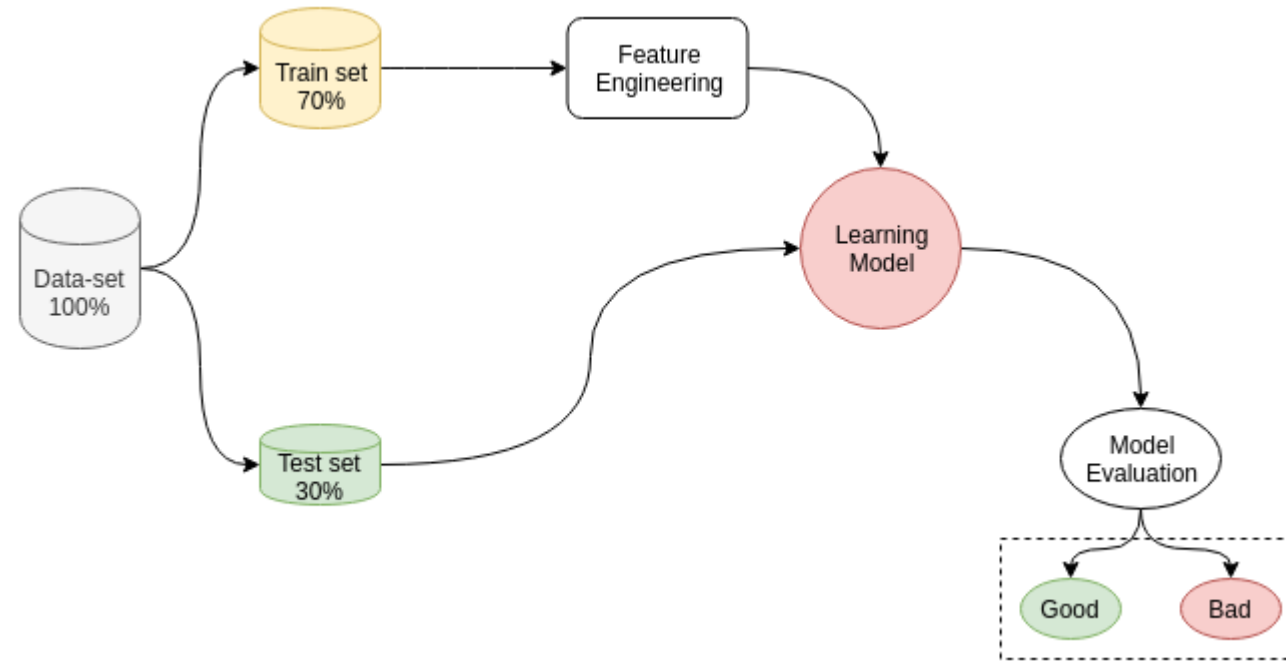
# Hiperparámetros

- Configuraciones que pueden controlar el comportamiento de un algoritmo
- Muy específicos a cada algoritmo
  - Get\_params()
- ¿Cómo lo hacemos?
  - Validación cruzada

# Validación cruzada



# Validación cruzada



```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3,
random_state = 42, shuffle = True, stratify = y)
```

# Técnica de Validación cruzada

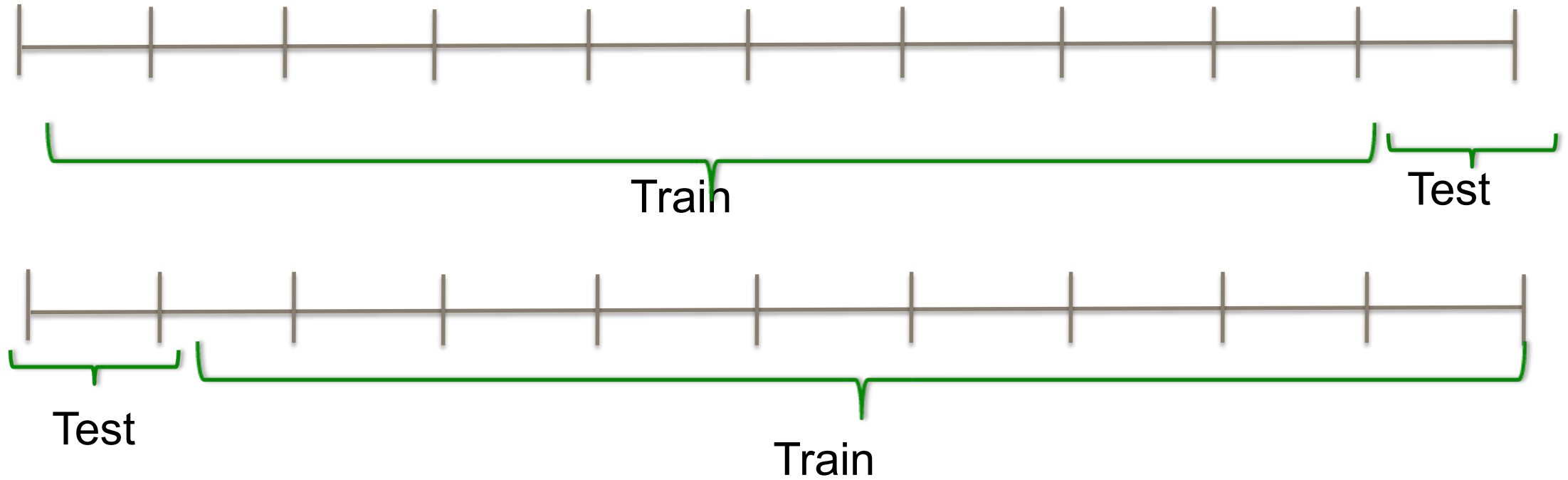
- K-Fold Cross Validation
- Leave P-out Cross Validation
- Leave One-out Cross Validation
- Repeated Random Sub-sampling Method
- Holdout Method

# K-fold Cross Validation

- La Validación cruzada es el modo más popular de evaluar un algoritmo en un conjunto de datos
- Necesitamos un conjunto de datos, un algoritmo, y una medida de evaluación de calidad del resultado. La medida puede ser el error cuadrado entre las predicciones y la verdad
- Divide los datos en 10 “pliegues” de aproximadamente el mismo tamaño
- Entrena el algoritmo en 9 de ellos, y calcula la medida de evaluación en el décimo.
- Repetimos este proceso 10 veces, utilizando cada “pliegue” en turnos como test
- Registramos la media y la desviación estándar de la medida de evaluación sobre los 10 “pliegues”



# K-fold cross validation



# K-Fold Cross Validation

```
from sklearn.model_selection import cross_val_score  
print(cross_val_score(model, X, y, cv=5))
```

```
import numpy as np  
print(np.mean(cross_val_score(model, X, y, cv=5)))
```

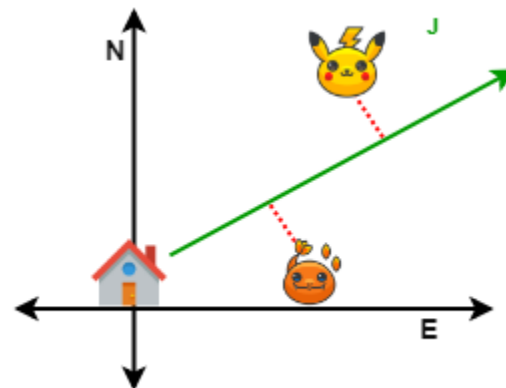
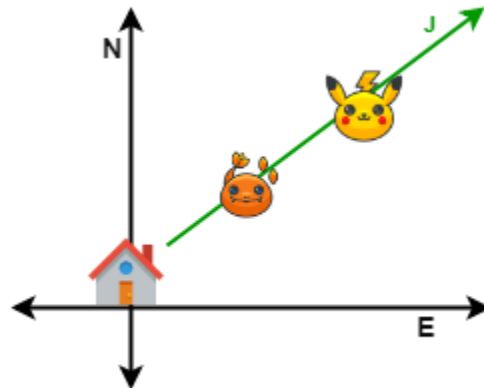
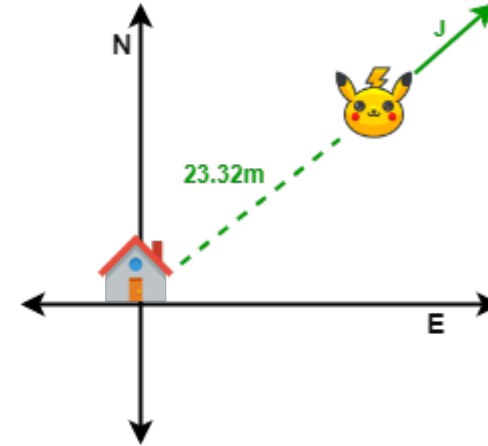
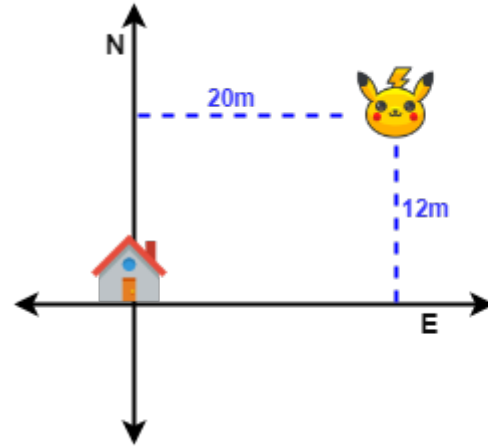
# Reducción de dimensionalidad



# Reducción de dimensionalidad

- Seleccionar las características más adecuadas
- ¿Cómo?
  - Selección de variables (ya lo hemos visto)
  - Análisis de componentes principales (PCA)
    - construir una dimensión (o variable) nueva a partir de la fusión de dos ya existentes.

# PCA





[www.solidq.com](http://www.solidq.com)

[info@solidq.com](mailto:info@solidq.com)