



# SODDISFAZIONE DEI CRITERI RICHIESTI

# ORIGINALITÀ E SIGNIFICATIVITÀ

- Abbiamo impiegato soluzioni «self-made» per l'estrazione di dati dal dataset con successiva manipolazione di questi ultimi per l'inserimento automatico di atomi all'interno della base di conoscenza.
- Abbiamo utilizzato la Grid Search per cercare la migliore combinazione degli iperparametri nei diversi algoritmi di regressione.

# COPERTURA

- Abbiamo coperto diversi argomenti del corso:
  - Estrazione di features tramite analisi del dataset;
  - Uso di diversi algoritmi di regressione per effettuare la predizione;
  - Si sono applicati i modelli come Lasso, RidgeRegression ed ElasticNet in modo da utilizzare un regolarizzatore per penalizzare i modelli complessi;
  - Uso di una base di conoscenza ricavata dal dataset.

# COMPLESSITÀ

- La complessità del dataset ha richiesto un'attenta fase di analisi per estrarre le features più rilevanti ed individuare eventuali features nascoste.

# GENERALITÀ

- Possibilità di estendere il progetto ad altri domini (ad esempio altre risorse differenti dalle auto);
- Possibilità di aggiungere ulteriore conoscenza ampliando i task del progetto, come ad esempio la possibilità di sfruttare la base di conoscenza a fini diagnostici.

# QUALITÀ DELLA VALUTAZIONE

- Abbiamo addestrato diversi regressori per individuare quello migliore utilizzando come strumento di comparazione le seguenti metriche di valutazione:
  - MSLE (Mean Squared Log Error);
  - R2 Score.
- Abbiamo scelto di adottare il Mean Squared Log Error poiché la feature target è una variabile continua con valori di dominio dell'ordine delle migliaia.

The background is a blue gradient with decorative white circuit-like lines in the corners. The lines consist of straight segments and small circles, resembling a stylized electronic circuit.

# GRAZIE PER L'ATTENZIONE

Alessandro Papeo

Michele Stelluti

Vincenzo Susso

Giuseppe Tanzi