# SARA ALLALI

(714)-768-7208   sallali@chapman.edu   Sara Allali | LinkedIn   Sallali02 (Sara Allali) (github.com)

## PROFESSIONAL SUMMARY

Data Science graduate skilled in machine learning, deep learning, NLP, and statistical modeling, with experience building predictive models, CNNs, and NLP pipelines to extract insights and optimize decision-making. Proficient in Python, SQL, R, and cloud analytics tools with a track record of delivering scalable, data-driven solutions. Passionate about applying AI and advanced analytics to solve complex problems and drive measurable impact.

## EDUCATION

**Chapman University, California -** M.S Computational and Data Sciences                Graduated: May 2025
**Chapman University, California** - B.A Psychology, Minor Analytics                Graduated: May 2024
Relevant Coursework: Statistical Machine Learning, Deep Learning, Natural Language Processing, Data Structures and Algorithms, Statistical Models in Business Analytics, Mathematical Modeling, Data Mining

## TECHNICAL SKILLS

**Programming & ML:** Python, SQL, R, Java, C++, Scikit-Learn, TensorFlow, PyTorch,
**Data Visualization:** Power BI, Tableau, Excel, Matplotlib, Seaborn
**Cloud & Analytics Platforms:** AWS, Snowflake, Databricks, Salesforce
**ETL & Workflow:** Alteryx, Jupyter Notebook, Google Colab
**NLP & AI Tools:** Hugging Face, spaCy, NLTK, Transformers
**Research Tools:** Qualtrics, MS Forms, Survey Design & Analysis

## WORK EXPERIENCE

**AI Research Fellow |** Klarbook, Remote
July 2025 - Present
- **Collaborated** on a national AI research study analyzing structured and unstructured data to uncover industry trends in automation and AI.
- **Coordinated data collection** through structured outreach and survey methods, ensuring accurate and organized real-world data.
- **Contributed insights** to a published benchmark report and gained hands-on experience with modern AI tools and methodologies.

**Data and Systems Analyst Assistant |** Chapman University, Orange, CA
October 2023 - May 2025
- **Cleaned and analyzed** 100k+ student records to ensure data accuracy and deliver actionable insights for institutional decision-making.
- **Extracted and managed** data from multiple platforms to generate insights aligned with leadership needs.
- **Designed and developed** interactive dashboards, enhancing decision-making for leadership
- **Assisted in survey development** for campus-wide student research, enabling data collection for key institutional insights.
- **Built dashboards to track student outcomes**, helping career services evaluate post-graduation success and program effectiveness.
- **Developed** logistic regression models to identify key predictors of graduate employment outcomes, improving career services' targeted interventions and resource allocation.
- **Collaborated** with the Data and Systems Analyst to ensure data integrity and project alignment.
- **Reviewed and validated** data before publishing reports to maintain consistency and accuracy.

- **Supported ad hoc projects**, showing adaptability and problem-solving in a data-driven setting.

## RELEVANT PROJECTS

**Eye Disease Classification Using Deep Learning** | Python, TensorFlow, [Project Link](#)
- Built a CNN to classify retinal images into four disease categories (~1K images/class), achieving high accuracy for automated diagnosis.

**Predictive Model for Neonatal Sepsis Detection** | SQL, R, SMOTE, SHAP, [Project Link](#)
- Built ML model to detect neonatal sepsis (ROC 0.93, 99.8% sensitivity) with SHAP-based interpretability, improving early diagnosis in clinical workflows.

**Song Generation with NLP and Deep Learning** | Python, NLTK, TensorFlow, LSTM, [Project Link](#)
- Trained an LSTM-based NLP model on 25K+ songs to generate new lyrics, showcasing AI's creative potential.

**Multimodal AI for Radiology: Integrating Image and Text for Enhanced Captioning and VQA in Medical Imaging** | Python, NLP, PyTorch, transformers, [Project Link](#)
- Integrated X-ray images and radiology text reports using CLIP/BLIP for automated captioning & Visual Question Answering (VQA).

**Automated Video Content Summarization** | Python, spaCy, NLTK, Transformers, BART, [Project Link](#)
- Built an NLP pipeline using BART/T5 transformers to summarize YouTube subtitle transcripts into concise bullet points.

## LEADERSHIP & INVOLVEMENT

- **Career Fair Support, Chapman University** (Spring 2024, Fall 2024, Spring 2025)
  Assisted in planning and executing one of the university's largest career fairs, enhancing employer-student engagement.
- **Giving Day Photoshoot Participant, Chapman University** (Annual Giving Programs)
  Supported marketing initiatives that promoted fundraising and campus engagement.
- **Club & Organizations**: Member of Girls Who Code (GWC) and Data Analytics Association

## CERTIFICATIONS & HONORS

- **Certifications & Training:** Data or Specimens Only Research (CITI Program), Tableau Desktop: Speed & Performance Optimization, The Complete SQL Bootcamp: Go From Zero to Hero
- **Awards & Recognition**: Provost's List (2021-2024) | Honor's Magna Cum Laude (2024)