

# earn data science by building

# DATA SCIENCE ACADEMY: DECISION MAKING USING DATA



## ENTERPRISE TRAINING TESTIMONIALS



### **Alim Kidar Hanif** Tokopedia

During my first week, I was not really able to cope with the lesson like statistics, but with the help of Algortima's mentors, it became somewhat easy to understand statistics.



### Rhendy Akbar Hilman

Kementerian Kesehatan RI

The lecturers are very experts in this field, able to master the materials and liven up the classroom atmosphere.



### **Enos Eben Ezer** SKK Migas

Here, we learn that our data can be utilized optimally to provide deeper insights before making decisions.



### Asep Taufik Hidayat Telkomsigma

So, my key takeaway from the course is the utilization of the data visualization with R, library in R, step-by-step preprocessing, and then we get to know how to use a lot of methods in data science and the data criteria suitable for our case.



### Hafiz Sudin Amin BPK RI

The instructor team of Algoritma are absolutely capable, they not just master the theory, but they also have practical experiences so that we are taught with materials that are more relevant to the field that we are in.



# STUDENT TESTIMONIALS



### **Ezra** Cohort Jupiter

The lifetime learning benefit provided is really beneficial and can keep us up-to-date with the current data science trends.



### Albers Uzila Dwiyanda Cohort Midas

Algoritma encourages students to continue learning independently by providing a project called "learning by doing."



### **Risal Andika** Cohort Hydra

The material provided is not just practical, but also accentuating the business side so that it became relevant to the real-life business case.



### Ronny Goei Cohort Kappa

Algoritma is active for career support from making the final project, selecting topics, helping to build our foundation, and presenting materials.



### Alvernia Eka Poetry Cohort Deragon

Fortunately, because I joined Algoritma Academy and with the provision of the community and teaching assistant, so whenever I need them, I just need to send an email or message.



### Naufal Abdila Cohort Midas

After taking Algoritma courses, I became more insightful on how data science is used and how we can leverage the data to make a business decision. Not only for predicting, but we can also create historical data to analyze the trends we haven't seen before.



# **REAL BUSINESS CASES**

Digital transformation has accelerated the amount of data generated by businesses and organizations around the world. Companies that can transform their data into actionable insights will thrive and gain a competitive advantage.

The increasing awareness of the importance of data-driven decisions has spiked up the demand for workforces with solid business acumen, a deep understanding of statistics, and the ability to work with large amounts of data.

Algoritma Data Science Academy is a 4 months program that will help you to develop your programming & statistics skills as you build data science projects modeled after real-life business cases, one at a time.

Algoritma Academy consists of three specializations:



Gain proficiency in creating data products and dynamic dashboards to extract insights using R Programming language.



Understand the principles and application of machine learning algorithms to solve business problems.



Obtain full proficiency in data analysis using Python and SQL.





# **LEARNING PROJECTS**



## PERSONALIZED LEARNING

Every student have different learning objectives, abilities, and preferences. We will personalize your learning experience with:



### **Interactive Learning**

An experienced Instructor leads our online and onsite classes. Our online program is delivered via Zoom, and you will feel an interactive experience as if you are present in a physical classroom.



### **In-Class Mentoring**

Our Teaching Assistant will provide in-class help and support when you encounter difficulties following along.



#### **Group Mentoring**

Are you feeling left behind or need help with your homework or projects? You can join the weekly mentoring session to help you get unstuck.



# Every Student is Unique.



Course Name	Objective	Knowledge	Description	Key Competencies
Python for Data Analysts (P4DA). I2 HoursPython environments and IDE.Python Basic Programming.Introduction to Pandas library.	Python environments and IDE.	<ul> <li>Understanding the practice of using package &amp; environment managers in python projects.</li> <li>Proficiency in writing Python codes in Jupyter Notebook.</li> </ul>	<ul> <li>Package is the mandatory part to help in analyze, explore, and visualize data. In addition, learning packages and environment manager will help the student to build and maintain projects easily.</li> <li>Jupyter Notebook is a platform that allows the student not only to write the script but also to create documentation. Understanding Jupyter Notebook will help the student to present project documentation as a product report.</li> </ul>	<ul> <li>Knowing what and why virtual environment.</li> <li>Differentiate between Python Script and Python Notebook.</li> <li>Understanding Python Notebook document.</li> </ul>
	Python Basic Programming.	Understanding Python fundamentals.	Python is one of the programming languages used for data analysis. Therefore, learning Python fundamentals is essential before data processing.	<ul> <li>Knowing basic Python syntax.</li> <li>Differentiate between referencing and copying in Python.</li> </ul>
	Introduction to Pandas library.	<ul> <li>Understanding the use of Pandas library as data analysis tools in Python.</li> <li>Performing exploratory data analysis in Pandas.</li> <li>Indexing and subsetting with Pandas.</li> </ul>	<ul> <li>Pandas is one of the powerful libraries used by Python to help in the data analysis process. Understanding Pandas library will make student easier in wrangle and analyze data.</li> <li>Exploratory data analysis (EDA) is an essential part of analyzing patterns of the data. After understanding the Pandas library, students will have the ability to recognize data patterns.</li> </ul>	<ul> <li>Knowing data structures in Pandas.</li> <li>Knowing data types in Pandas.</li> <li>Read and explore data.</li> <li>Understanding basic descriptive statistics in .describe() method.</li> <li>Able to perform indexing and subsetting operations with Pandas.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
Exploratory Data Analysis (EDA) 12 Hours	Working with date, time & categorical data.	<ul> <li>Familiarity with DateTime format in the data analysis process.</li> <li>Knowing alternative solutions to handle data type conversion in more extensive data.</li> </ul>	<ul> <li>Ability in understanding DateTime format allows the student to create manipulation in DateTime data type for further analysis.</li> <li>Data type conversion is essential to make the data manipulation more effective and efficient.</li> </ul>	<ul> <li>Knowing what distinguishes the DateTime object from other data types.</li> <li>Knowing how to extract relevant time information from a DateTime object.</li> <li>Knowing how to combine pandas slicing method and its .apply() function.</li> </ul>
	Why and what: exploratory data analysis.	Understanding the use of contingency tables and how to create it in Pandas.	Understanding in producing contingency tables allows students to present information about frequency or data aggregation.	<ul> <li>Knowing what, why, and how to create a frequency table.</li> <li>Knowing what, why, and how to create an aggregation table.</li> </ul>
	Treating missing values and duplicates in data.	<ul> <li>Understanding missing values in data and how to handle them.</li> <li>Understanding duplicates data and how to remove them.</li> </ul>	Handling missing and duplicated values are necessary to create better analysis to answer business problems.	<ul> <li>Knowing how Pandas denote missing data.</li> <li>Able to observe and impute missing values in data.</li> <li>Able to monitor and handle duplicates.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
Data Wrangling & Visualization (DWV) 12 Hours	Python reproducible environment.	Understanding how to handle dependencies in a Python project.	Learning and understanding how to create a reproducible environment are mandatory since they will help standardize requirements between team members.	Able to import and export Python dependencies.
	Understanding multi-index dataframe.	Able to identify/differentiate multi-index dataframe.	A multi-index dataframe is a high-dimensional dataframe with two or more indexes in rows or columns. Ability in understanding structure multi- index dataframe is essential since multi-index dataframe have different treatment from single index dataframe.	<ul> <li>Able to fetch data from various data sources.</li> <li>Able to work with multi-index dataframe.</li> </ul>
	Data wrangling and reshaping.	<ul> <li>Able to perform data reshaping with Pandas.</li> <li>Able to perform group by aggregation.</li> </ul>	<ul> <li>Data reshaping for data preparation and visualization needs.</li> <li>Group by aggregation is essential to present hidden information within the data.</li> </ul>	<ul> <li>Knowing why and how to perform data reshaping with Pandas.</li> <li>Able to aggregate data using group by operation.</li> </ul>
	Data Visualization.	Able to create data visualization using Pandas for exploratory data analysis.	Showing data patterns from the exploratory results using visualization method, helping the student to understand insight from the data easily.	<ul> <li>Knowing how to interpret &amp; create a plot in Pandas.</li> <li>Using boxplot to visualize statistical distribution in data.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
SQL Query 12 Hours	Working with SQL database.	<ul> <li>Able to work with SQL database.</li> <li>Understanding in under and overfetching.</li> </ul>	<ul> <li>By understanding SQL and Python implementation, students will have the ability to fetch data using SQL then analyze and wrangle the data using Python.</li> <li>Under-fetching is a condition where the data loaded is less than required, while over-fetching is the opposite. Both of them have a significant impact on data preparation. Under-fetching leads to non-insightful information, while over-fetching leads to inflated cost.</li> </ul>	<ul> <li>Connecting SQL database.</li> <li>Writing queries.</li> <li>Perform SQL aggregation.</li> </ul>
Capstone Project - Data Analytics <b>3 Hours</b>	End-to-end data project.	Understanding in end-to-end project analysis.	Students will perform data analysis using Python programming language by building mini projects based on their learning.	Knowing several alternatives to publish analysis reports.



Course Name	Objective	Knowledge	Description	Key Competencies
Introduction to Machine Learning I (IML1) 12 Hours	Basic Principles of Machine Learning.	<ul> <li>Understanding of the basic principles of machine learning and its applications.</li> <li>Able to decide predictor-target variables and preprocess data before model machine learning building.</li> <li>Understanding the types of common machine learning cases—classification and regression.</li> </ul>	<ul> <li>This module covers fundamental concepts of machine learning using the scikit-learn library. Students will learn data representation (predictor and target variables), loading, and preprocessing, as well as classification and regression cases in machine learning.</li> </ul>	<ul> <li>Representation of Data in the Scikit-learn.</li> <li>Load Data &amp; Data preprocessing.</li> <li>Predictive Analysis with ML: Classification &amp; Regression.</li> </ul>
	Machine Learning Workflow: Support Vector Machine (SVM).	<ul> <li>Understanding the concept of Support Vector Machine and its hyperparameters.</li> <li>Knowing the general machine learning workflow.</li> </ul>	• Support Vector Machine (SVM) is a family of extremely powerful models that can be used in classification and regression problems. Students will learn about SVM concepts, including the maximum margin classifier, and the concept of its hyperparameters	<ul> <li>Introduction to Support Vector Machine (SVM).</li> <li>Hyperparameters SVM.</li> </ul>
	Implementation SVM in Classification and Regression.	<ul> <li>Able to create classification ("Support Vector Classifier") model.</li> <li>Able to create regression ("Support Vector Regressor") model.</li> </ul>	<ul> <li>SVM can solve classification ("Support Vector Classifier") and regression ("Support Vector Regressor") problems and can be extended to model non-linearity in the data. Even though the workflow between classification and regression is almost the same, the approach from model building to evaluation will be different.</li> </ul>	<ul><li>SVM in Classification.</li><li>SVM in Regression.</li></ul>



Course Name	Objective	Knowledge	Description	Key Competencies
	Model Improvement Technique.	• Proficiency in evaluating models with the right evaluation metrics—accuracy for classification models and mean absolute error (MAE) for regression model.	• Students will improve model performance by tuning its hyperparameters and how to measure model performance. Model evaluation is necessary because we can know the model is good or not (can predict correctly or not) from this step.	<ul> <li>Model Evaluation.</li> <li>Tuning Hyperparameters.</li> </ul>
Introduction to Machine Learning II (IML2) 9 Hours	Ensemble Machine Learning using XGBoost.	• Understanding ensemble machine learning using XGBoost fundamentals.	<ul> <li>Ensemble machine learning is essential to know because its more potential to get a model that generalizes well to unseen data (low bias and low variance) than a single machine learning model.</li> <li>One of the methods of ensemble learning is boosting. For this course, students will be looking at a specific algorithm that uses the boosting method and that is the XGBoost algorithm.</li> </ul>	<ul> <li>Introduction to Ensemble Learning.</li> <li>Types of Ensemble Learning: Boosting.</li> <li>A Short Introduction to Boosting.</li> <li>Concept of XGBoost and Its Hyperparameters.</li> </ul>
	Implementation XGBoost in Classification and Regression.	<ul> <li>Able to create classification ("XGBClassifier") model.</li> <li>Able to create regression ("XGBRegressor") model.</li> </ul>	<ul> <li>Covering up how to use XGBoost for classification and regression tasks, including hands-on activities and examples.</li> </ul>	<ul><li>XGBoost in Classification.</li><li>XGBoost in Regression.</li></ul>
	Tuning Hyperparameter Optimizer.	<ul> <li>Proficiency in tuning models more efficiently and effectively to get the best model performance using GridSearchCV and RandomSearchCV.</li> </ul>	<ul> <li>GridSearchCV and RandomSearchCV are techniques for hyperparameter optimization and how to use them to improve machine learning model performance.</li> </ul>	<ul><li>GridSearchCV.</li><li>RandomSearchCV.</li></ul>



Course Name	Objective	Knowledge	Description	Key Competencies
Programming for Data Science (P4DS) 9 Hours	Programming Language and Tools in R.	<ul> <li>Building reproducible data analysis reports using Rstudio.</li> <li>Proficiency in using the R programming language.</li> </ul>	RStudio is the main graphical user interface (GUI) for the R programming language. Therefore, learning the environment of RStudio will help the student effectively use R and RStudio as primary tools for a data science project.	<ul> <li>Rstudio Environment.</li> <li>Differentiate between R Markdown and R Script.</li> <li>R Syntax.</li> <li>Data types in R.</li> <li>Data structures in R.</li> </ul>
	Data Manipulation.	Data manipulation is often required to perform analysis. Therefore, learning the basic data manipulation technique is essential.	Able to manipulate data in R.	<ul> <li>Read and extracting data.</li> <li>Data cleansing methods.</li> <li>Data transformation methods.</li> </ul>
	Statistical Computing.	Able to use R as a data science tool.	Data scientists are expected to create a product, be it an analytics report, machine learning models, or a script for automating the process. Therefore, learning how to present a report from Rmarkdown or building a custom function via R script is an essential part of the data science workflow.	<ul> <li>Knowing project workflow with R.</li> <li>Making reproducible data science with R.</li> </ul>
Practical Statistics (PS) 6 Hours	Descriptive Statistics.	Summarise data and interpret correlation. Able to analyze data distribution and identifying outliers.	<ul> <li>Descriptive statistics is a quick way to summarise and gain insights from data. Learning descriptive statistics is also essential to build accurate and effective data visualization.</li> <li>Data distribution will give more insight into the overall data than a single value such as mean and</li> </ul>	<ul> <li>Measurement of central tendency data.</li> <li>Measurement of spread.</li> <li>Measurement of data relation.</li> <li>Benchmark outlier of data.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
			<ul> <li>median. Detecting outlier or extreme is essential since an outlier can be a sign of anomaly, rare event, error, or anything that may affect the result of the analysis.</li> <li>Normal distribution is commonly found in data and is a basis for various statistical testing and machine learning algorithm.</li> </ul>	• Parameter of normal distribution.
	Statistical Inference.	<ul> <li>Able to calculate probability.</li> <li>Able to perform hypothesis testing analysis.</li> </ul>	<ul> <li>The probability distribution theory is essential and one of the main principles behind various statistics hypothesis testing.</li> <li>Hypothesis testing is essential in many business processes, especially where we want to do some experiments and analyze the result.</li> </ul>	<ul> <li>Probability testing.</li> <li>Confident of a sample.</li> <li>Statistical testing.</li> <li>Hypothesis testing.</li> </ul>
Data Visualization in R (DVR) 12 Hours	Plotting Essentials.	Able to create a statistical plot in R.	A data visualization, first and foremost, has to convey the data accurately. Therefore, students will learn the basics of data visualization, including graphic elements, what type of charts are appropriate for numerical data, what type of charts are suitable for categorical data, etc.	<ul> <li>Understanding built-in plotting functionalities in R.</li> <li>Understanding how to customize plot with built-in function in R.</li> </ul>
	Plotting With ggplot2.	<ul><li> Able to create a plot with ggplot2.</li><li> Able to create a publication-ready plot.</li></ul>	• ggplot2 is the best package for visualization in R and has been recognized by some major companies, including Airbnb, BBC, Google, etc. By learning ggplot2, students will able to build a practical and aesthetically pleasing visualization.	<ul> <li>Understand grammar of graphics system.</li> <li>Understand geometries.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
			<ul> <li>Good visual presentations tend to enhance the message of the visualization. It will help convince clients and people about the argument and insight that the data scientist/analyst finds.</li> </ul>	• Understand how to create a multi- dimensional faceting plot.
	Introduction to Leaflet.	Able to visualize geospatial data.	Building an effective visualization with map and geospatial data will help the student to present more insights.	Understand the basic concept of leaflet.
Interactive Plotting and Dynamic Dashboards (IPDD) <b>12 Hours</b>	Data Wrangling with dplyr, and Plotting with Plotly.	<ul> <li>Able to process data with dplyr.</li> <li>Able to create interactive plot using plotly</li> </ul>	<ul> <li>Data wrangling or data transformation is an essential step in data science workflow before doing any analysis or visualization. The dplyr package is a great tool that will help the student to do data wrangling.</li> <li>Interactive plot can give more information compared to static visualization. People can hover or click the plot to access more information from each data point.</li> </ul>	<ul> <li>Understanding how dplyr is works.</li> <li>Common function to do data preparation with dplyr.</li> </ul>
	Building Interactive Document.	Able to create an interactive document with flexdashboard.	A dashboard is essential to present many insights and information on a single page. Flexdashboard is a package from R that will allow students to build interactive plots and design a single-page dashboard.	Understanding how to create a dashboard with flexdashboard.
	Introduction to Shiny.	<ul> <li>Able to create a dashboard using Shiny.</li> <li>Able to publish Shiny application online.</li> </ul>	• A web dashboard is a great tool to present real-time information for management or other departments. Sometimes the information will be divided into multiple tabs/pages. Shiny is an excellent package from R to build a web dashboard and have the same features as Tableau or PowerBI.	<ul> <li>Understand how to build an interactive dashboard.</li> <li>Understand Shiny dashboard components.</li> <li>Understand hosting and deployment.</li> </ul>



Course Name	Objective	Knowledge	Description	Key Competencies
			• A dashboard should be accessible for people who need them. Deploying the shiny dashboard to the online web will make other people has access to the information and visualization in real-time.	
Capstone Project - Data Visualization <b>3 Hours</b>	End-to-End Data Visualization Project.	Able to analyze data and build an interactive data visualization dashboard.	Data scientists should find insight and recommendations from data and communicate the result on a practical and interactive web dashboard.	<ul> <li>Prepare data for visualization.</li> <li>Build a data visualization dashboard.</li> <li>Communicate insights from the dashboard.</li> </ul>



### MODULES & LEARNING OUTCOMES MACHINE LEARNING SPECIALIZATION (90 HOURS)

Course Name	Objective	Knowledge	Description	Key Competencies
Regression Models (RM) 12 Hours	Linear Regression Models and the Key Terminologies.	<ul> <li>Data preparation.</li> <li>Create and interpret linear regression model.</li> <li>Evaluate the linear regression model.</li> </ul>	<ul> <li>As data scientists, we must create and explain the machine learning models that have been created. In addition, model interpretation is needed to know each predictor's effect on the predicted results.</li> <li>After making a regression model, we want to get the best model performance to predict our new data set. To determine the best model, we must evaluate it using its model performance.</li> </ul>	<ul> <li>Linear regression concept.</li> <li>Feature selection.</li> <li>Evaluating linear regression.</li> <li>Linear regression assumptions.</li> </ul>
Classification in Machine Learning I (CML 1) <b>12 Hours</b>	Build Logistic Regression and K - Nearest Neighbors Algorithm.	Able to choose the suitable classification method to solve business problems.	There are various machine learning models ready to be deployed. However, understanding the most basic and simple algorithms is essential. Logistic regression is sufficient to solve most of the problems while also comes with the benefit of interpretability for research. K-NN is a simple algorithm that can deal with multiple numeric predictors.	<ul> <li>Logistic Regression Concept.</li> <li>K - Nearest Neighbors Concept.</li> </ul>
	Solving Classification Problem.	<ul> <li>Create classification models.</li> <li>Evaluating classifiers.</li> </ul>	As data scientists, we must create models to predict classes / categorical data, explain the model, and evaluate the model's performance.	<ul> <li>Logistic regression model.</li> <li>K - nearest neighbors model.</li> <li>Using classification model.</li> <li>Model evaluation.</li> </ul>
Classification in Machine Learning II (CML 2) <b>12 Hours</b>	Understand the Practical and Theoretical Aspects of Naive Bayes, Decision Tree, and Random Forest Model.	Able to choose a suitable classification method.	After students understand the basic machine learning algorithm, it will be essential to learn about more advanced classification models. Understanding the characteristics, advantages, and disadvantages of different models will help students to know which model to use to solve the business problem effectively.	<ul> <li>Naive Bayes concept.</li> <li>Decision Tree concept.</li> <li>Random Forest concept.</li> <li>Random Forest model evaluation.</li> </ul>



### MODULES & LEARNING OUTCOMES MACHINE LEARNING SPECIALIZATION (90 HOURS)

Course Name	Objective	Knowledge	Description	Key Competencies
	Text Mining in R.	Able to do data preprocessing from raw text	Many data may not come in a structured (tabular) format. Text is an example of widely used unstructured data. It is essential for a data scientist to be able to process text data to gain insight and create a machine learning model.	<ul><li>Text cleansing.</li><li>Implementation of Naive Bayes for text classification.</li></ul>
Unsupervised Machine Learning (UL) <b>12 Hours</b>	Practical and Theoretical Aspect of PCA and K-Means.	<ul> <li>Determine appropriate PCs for dimensionality-reduction and reconstruct the data using the corresponding PCs.</li> <li>Determine reasonable k and evaluating clustering result.</li> </ul>	<ul> <li>High dimensional data contains more information but may also include noise and needs higher computational power to solve. Using appropriate PCs allows us to reduce data dimension while retaining as much information as possible, resulting in a decent model with less effort.</li> <li>Understanding K-Means and clustering will allow students to make meaningful groups (cluster) of data and interpret for each cluster to obtain business insight which will help in business decision making.</li> </ul>	<ul> <li>The basic concept of PCA.</li> <li>The basic concept of K-Means.</li> <li>Evaluating clustering result.</li> </ul>
	Apply Biplot for Visualization and Draw Insight.	Create biplot using desired PCs. Understanding of biplot.	Visualization of multi-dimensional data is highly complex. Biplot helps us visualize and understand multi-dimensional data, such as detecting outlier of the overall data, correlation between features, data distribution, etc.	<ul> <li>Qualitative and quantitative variables.</li> <li>Biplot visualization.</li> <li>Biplot's feature.</li> </ul>
Time Series & Forecasting (TSF) <b>12 Hours</b>	Understand the Concept of Time Series and Components of the Decomposition.	<ul> <li>Able to create time series object.</li> <li>Able to extract components trend, seasonal, and error from a time series object.</li> </ul>	<ul> <li>A time-series object is an object that contains the value itself and time factor. It is needed to make a time series analysis.</li> <li>Trend, seasonal, and error are components that exist in many time series objects. These components will help us to determine the right time series model to train.</li> </ul>	<ul><li>The basic concept of time series.</li><li>Time series decomposition.</li></ul>



## MODULES & LEARNING OUTCOMES MACHINE LEARNING SPECIALIZATION (90 HOURS)

Course Name	Objective	Knowledge	Description	Key Competencies
	Application of Time Series Model	<ul> <li>Implementation of time series model for forecastings.</li> <li>Evaluate the time series model.</li> </ul>	<ul> <li>Many time series model has a requirement in the time series component, so we can't randomly use the model. Understanding the requirement of each time series model will help us to determine a suitable model.</li> <li>Model evaluation is used to evaluate the model created to forecast our future data.</li> </ul>	<ul> <li>Forecasting models.</li> <li>The linear relationship between lagged values of a time series.</li> <li>Model evaluation.</li> <li>Assumption of time series model.</li> </ul>
Neural Network & Deep Learning (NN&DL) 12 Hours	Application of Neural Network.	<ul> <li>Knowing the correct activation function and cost function when building a neural network architecture.</li> <li>Neural network application.</li> <li>Able to use Keras framework to build a neural network model.</li> </ul>	<ul> <li>The activation function and the cost function are essential things in neural network architecture. It will determine your model's prediction accuracy.</li> <li>Keras is one of the deep learning frameworks. This framework will help to build a neural network model easier.</li> </ul>	<ul> <li>Neural network concept.</li> <li>Data preparation.</li> <li>Build neural network model.</li> <li>Evaluating neural network model.</li> </ul>
Capstone Machine Learning <b>3 Hours</b>	End-to-end Machine Learning Workflow.	Able to build end-to-end machine learning workflow.	Student will be given a series of cases to solve with machine learning model, such as predicting the number of customers in the next week (time series), building image recognition model (classification), or predicting the strength of a concrete mixture based on the materials composition (regression).	<ul> <li>Knowing how to do data preprocessing.</li> <li>Knowing how to build a supervised machine learning model.</li> </ul>



# CERTIFICATES

Upon successful completion of the academy program, participants will be awarded a digital certificate of completion and grade transcript by Algoritma Data Science School.





### **Our Vision:**

"To accelerate data science adoption across the region and provide employment opportunities."



# ALGORITMA DATA SCIENCE SCHOOL

We strive to solve the issue of an under-equipped modern workforce in facing the collective problem of the contemporary digital era. We help equip every professional with data science skills, such as data visualization, regression, data modeling, machine learning, and statistical programming, that will prepare them for today and tomorrow's employment.

Algoritma provides data science education by training individuals, corporates and provides employment opportunities for our graduates to various industries. We also provide consulting services to maximize value creation through data.

# THE 1ST TO BE CERTIFIED IN ASIA

RStudio's Instructor training and certification program is a rigorous program that applies modern evidence-based teaching practices to teach data science using R and RStudio. Our Instructors at Algoritma are <u>RStudio's Education official Partners</u> and the first in Asia to be certified.





 $\mathcal{A}$ 

ជ

2000 +

500 +

100 +

2

**Enterprise Clients** 

In-House Training









(in) Algoritma Data Science School

www.algorit.ma

