

AI Powered Predictive Analytics and Big Data Analysis for Business

Kohei Arai, Saga University

1

Course Description

- This course equips students with the knowledge and practical skills to leverage AI-driven predictive analytics and big data technologies for solving real-world business problems.
- Students will learn to preprocess and analyze large datasets, apply machine learning and AI models, and translate insights into actionable business strategies across domains such as marketing, finance, and operations¹²³⁴⁵.

2

Learning Outcomes

- By the end of the course, students will be able to:
 - Understand the role of **AI, machine learning, and big data in modern business decision-making**⁴⁵.
 - **Prepare, clean, and transform large datasets** for analysis¹²³.
 - Apply predictive analytics and **AI algorithms (regression, classification, clustering, deep learning)** to business data¹²⁶³.
 - Evaluate and interpret model performance using appropriate metrics and validation techniques¹⁶⁵.
 - Implement and deploy scalable AI solutions using big data platforms (e.g., Hadoop, Spark, cloud services)³.
 - Communicate analytics results and recommendations to stakeholders using visualizations and business reports²⁶⁵.

3

Course Modules & Topics

Module	Topics	Tools/Techniques
1. Introduction & Foundations	- Big Data definitions and taxonomy - Predictive analytics overview - CRISP-DM methodology - Value extraction from data	KNIME, Python, Power BI ¹⁶
2. Data Preparation & Processing	- Data cleaning, transformation, ETL - Handling structured & unstructured data - Data preprocessing at scale (cloud, Spark)	Python (pandas), Hadoop, Spark ¹²³
3. Statistical & Exploratory Data Analysis	- Descriptive statistics - Data visualization for business - Exploratory data analysis	Power BI, Tableau, matplotlib ²⁶⁵
4. Predictive Analytics Techniques	- Regression (linear, logistic) - Classification (decision trees, SVM, k-NN, Naive Bayes) - Clustering (k-means, hierarchical) - Time-series forecasting	scikit-learn, Power BI, AutoML

4

5. AI & Advanced Machine Learning	- Deep learning basics - Ensemble methods (bagging, boosting, random forest) - Natural Language Processing (NLP), chatbots - Reinforcement learning	TensorFlow/Keras, spaCy, cloud AI APIs ²³⁷
6. Model Evaluation & Deployment	- Model validation (cross-validation, ROC, confusion matrix) - Hyperparameter tuning - Deployment in real-time/big data environments	ML pipelines, cloud deployment (AWS/Azure) ¹²³
7. Business Applications & Case Studies	- Marketing analytics - Financial forecasting - Operations optimization - Ethics and data governance	Business case analysis, dashboards ⁴⁵
8. Capstone Project	- End-to-end AI-powered analytics project using real business data - Presentation of findings and recommendations	All above tools ⁶⁴

5

Assessment Methods

- **Practical assignments: Data cleaning, model building, and evaluation exercises.**
- **Group project: Real-world business problem solved using AI-powered predictive analytics.**
- **Presentations: Communication of insights and recommendations.**
- **Report by participants (Final exam or analytical paper):** Synthesis of theory and practical skills²⁴.
- **Final Presentations by 8 groups.**

6

Recommended Tools & Technologies

- Programming: **Python** (pandas, scikit-learn, TensorFlow/Keras, spaCy)²³⁷
- Big Data Platforms: Hadoop, Spark, cloud services (AWS, Azure)³
- Visualization: **Power BI**, Tableau⁶⁵
- Workflow Automation: **KNIME**
- **MySQL/MCP-RAG** ← **Dify**, **Cursor** (AI editor)

7

Suggested Readings

- *Applied Predictive Analytics* by Dean Abbott
- *Big Data in Practice* by Bernard Marr
- Online documentation for scikit-learn, TensorFlow, Power BI, and cloud AI services

8

Schedule

- (1) Introduction
- (2) One project: Group size ≤ 4 students, Checkpoints are as follows, Proposal: title and goal, Outline of approach, Implementation and Demonstration, Final Project Report by each student. Each group will have a short presentation and demonstration (15-20 minutes), Each group will provide a five-page document on the project; the responsibility and work of each student shall be described precisely.
- (3) Big Data definition, AI definition, Predictive Analysis, Decision Making Processes,
- (4) AI applications, Prediction Methods,

9

- (5) History of computers and AI, Inference, search, toy problem, expert system, General-purpose AI, specialized AI (strong AI/weak AI), Frame problem, symbol grounding problem, Changes occurring in society (big data, IoT (Internet of Things), AI, robots, the 4th industrial revolution: Industry 4.0, Society 5.0)
- (6) AI and Big Data as well as Predictive Analysis: Development of ICT (Information and Communication Technology), collection of big data, IoT, Big data platform, GAMAM, accumulation of big data, cloud service, Examples of big data utilization (human behavior log data, machine operation log data, SNS data, map data), Advanced AI (automatic driving, Siri, chatbot, etc.)

10

- (7) Data Science/Usefulness of Data: What is data science, Data scientist, Data engineering, Data analysis, Data distribution (histogram) and representative values (average, median, mode), Data variation (variance, standard deviation), Correlation, Data representation (line graph, bar graph, scatter diagram, heat map, etc.), Data aggregation (sum, average), Data sorting, ranking, Spreadsheet, CSV, JSON

11

Professor/s	Student Name	Student Surname	Country	Nationality	University	Signature
Prof. Dr. Kohei Arai Prof. Dr. Arben Pajaziti	1 Jade	Amponsah			UCL	
	2 Tilda	Tahiraj			Polytechnic University of Tirana	
	3 Tarig	Abdalla			Universiti Teknologi Malaysia	
	4 Elif	Südemer			Marmara University	
	5 Irem Pinar	Koroglu			Middle East Technical University	
	6 Varvara	Baradachova			SWPS University	
	7 Burak Demir	Izimek			Cukurova University	
	8 Arda	Simgir			Cukurova University	
	9 Maja	Brodowska			SWPS University	
	10 Britt	Kjorlien			Drake University	
	11 Samuela	Palmasiano			Politecnico di Bari	
	12 semiha	şeker			International University of sarajevo	
	13 Gabriel	Khouri			University of Michigan	
	14 Dominika	Pastuszka			Wyższa Szkoła Bankowa w Chorzowie	
	15 Mael Chan	Peng			INP TOULOUSE	
	17 Diellart	Muloll			University of Prishtina	
	18 Fiona	Bega			University of Prishtina	
	19 Rizaonë	Hyseni			University of Prishtina	
	20 Edonita	Gjyrevci			University of Prishtina	
	21 Iliride	Ferataj			University of Prishtina	
	22 Dea	Lajqi			University of Prishtina	
	23 Muhamed	Jakupi			University of Prishtina	
	24 Dea	Metaj			University of Prishtina	
	25 Lize	Sylla			University of Prishtina	
	26 Elda	Ferati			University of Prishtina	
	28 Elita	Jashari			University of Prishtina	
	29 Nol	Koxha			University of Prishtina	
	30 Anita	Ahmeti			University of Prishtina	
	31 Vera	Llugjaj			University of Prishtina	

12

Project candidates

- 1. Data preparation and cleansing: The first step in data analysis is to ensure high-quality data., Exploratory Data Analysis (EDA)
- 2. Web access and behavior analysis: Analytical methods for understanding user behavior
- 3. Customer understanding and segmentation: Analytical methods for gaining a deeper understanding of customers and grouping them
- 4. Customer value and satisfaction analysis: Analytical methods for optimizing relationships with customers

13

- 5. Marketing Effectiveness Measurement and Optimization: Analytical methods for measuring and improving the effectiveness of marketing activities
- 6. Lead Management/CRM Optimization: Analytical methods to optimize the process from acquiring prospects to converting them into customers
- 7. Predictive modeling and machine learning: Methods for future predictions and advanced data analysis
- 8. Corporate and market analysis: Methods for analyzing a company's internal and external environment

14

- 9. External environment and strategy analysis: Analytical methods related to market environment and strategy formulation
- 31 students=4 member by 8 groups

15

- From the site (URL: <http://teagis.jp.is.saga-u.ac.jp>), you can get the python codes for big data analysis for business.

新井康平, Kohei Arai

- [English Version Here](#)
- [Python code for AI powered predictive analysis and big data analysis is Here](#) 
- 佐賀大学名誉教授・客員研究員(Emeritus Professor Dr. Saga University)
- 西九州大学健康栄養学部客員教授・非常勤講師(Adjunct Prof. of Nishi-Kyushu University)
- 久留米工業大学AI応用研究所客員教授・非常勤講師(Adjunct Prof. of Kurume Institute of Technology)
- 日本学術会議(Science Council of Japan) 会員(COSPAR小委員会Sub Committee on COSPAR-Committee on Sapec Research)
- Brawijaya University (Adjunct Prof. of Brawijaya University)

16

8 Group Work

- 1G: Introduction & Foundations(KNIM, PowerBI, Python)
- 2G: Data Preparation (Cleansing, Cleaning, Python, Hadoop)
- 3G: EDA (PowerBI, matplotlib, Tableau)
- 4G: Predictive Analysis (Regression, classification, clustering, time-series analysis, scikit-learn, Power BI, AutoML)
- 5G: Deep Learning, NLP, Raienforcement Learning, chatbot, TensorFlow/Keras, spaCy, cloud AI APIs)
- 6G: Model Deployment (ML pipelines, cloud deployment)
- 7G: Business Application (Financial forecasting, marketing)
- 8G: MySQL/MCP-RAG (MCP host-client-servers, AI agent)

17

(8) Latest Trends in Data and AI Utilization:

- **AI/DS utilization examples** (Facebook, Microsoft search term ads, Amazon, Mercari: product recommendation system using collaborative filtering), **Dynamic Pricing**: Ticket pricing for planes, hotels, sports, etc. **Amazon Go**: Payment by facial recognition at stores without cash registers • **FourSquare**: Real-time advertising using GPS location information, **AI utilization of big data** (identification of product issues through natural language understanding of SNS data, IC cards, drive recorders (optimization of routes)), **Latest trends** (automatic driving, face recognition, police cases (crime prediction, etc.), satellite data AI, etc.), Investigation of AI/DS and predictive analysis utilization in local companies

18

(9) What is machine learning? Technology for data and AI utilization:

- **Supervised learning** (measured data and teacher data, training data, regression/classification) **linear regression**, **SVC**, **Unsupervised learning** (clustering, dimension removal), **Reinforcement learning** (automatic driving, robot control), **Pre-processing**, Missing data values, Normalization, Standardization

19

(10) AI and **Ethics**:

- **OECD 8 principles** for AI utilization, Pay attention to infringement of personal information (**privacy**), AI Social Principles, AI Development Principles, AI Utilization Principles, **AI risks**, Black box, Data bias problem, Vulnerability, **ELSI** (ethical, legal, and social issues) in social implementation, **Group discussion** on issues caused by data and AI.

20