DEGREE: MSc Data Analytics

Module: Fundamental of Data Analytics

Assignment Title: Machine Learning Meets Term Life Insurance: Targeting High-Value Customers

Learning Meet: Set exercise

Word Limit: 3000 words (+/- 300)

Weighting: 40%

Issue Date: 25/07/2024

Submission Date: 01/10/2024

Feedback Date: 22/10/2024

Plagiarism:

When submitting work for assessment, students should be aware of the InterActive/Canvas guidance and regulations in concerning plagiarism. All submissions should be your own, original work.

You must submit an electronic copy of your work. Your submission will be electronically checked.

Learner declaration
I certify that the work submitted for this assignment is my own and research sources are fully acknowledged. **Student signature:**   **Date:**

Harvard Referencing:

The Harvard Referencing System must be used. The Wikipedia, UKEssays.com or similar websites must **not** be used or referenced in your work.

Introduction

Learning Outcomes:

LO1. Demonstrate an understanding of basics of linear algebra, calculus, statistics and algorithms to structure
 basic solutions of business using data.

LO2. Use of statistics for business decision making and the ways to source relevant data to assist decision making & problem solving.

LO3. Solve problems using derivatives, transcendental functions, and integration, integrate the concepts of foundations of data analytics for problem solving.

Tasks (All tasks are equally weighted):

Objectives:

This data analytics assignment is designed for master's students seeking to solidify their foundational knowledge and practical skills in the field. This assignment will not only enhance your understanding of how data can be leveraged to solve business problems but also equip you with the critical thinking and analytical abilities necessary to thrive in today's data-driven world.

Scenario:

HashSysTech Insurance, a dynamic insurance provider known for its innovative approach, utilizes various outreach methods to promote its term life insurance products. Telemarketing campaigns have proven consistently effective in reaching potential customers. However, due to the high cost associated with these campaigns, HashSysTech seeks to optimize their resources.

You, a data analyst at HashSysTech, are tasked with developing a machine learning model to predict customer conversion for term life insurance. This model will play a pivotal role in Project Greenlight, an initiative designed to identify customers most likely to accept for term life insurance, allowing HashSysTech to strategically target them via telemarketing campaigns.

Dataset:

For this assignment, you will use the provided dataset, which can be accessed at the following link: HashSysTech Insurance Dataset (https://github.com/usmanakhtar/dataset/).

Task Phases:

Task 1: Data Exploration and Preparation (LO1, LO2, LO3)

1.1 Data Exploration

- **Objective**: Use descriptive statistics to understand the distribution of features and their potential relationships with the target variable (term life insurance conversion).
- **Instructions**:
  - o Calculate measures of central tendency (mean, median) and dispersion (standard deviation) for numerical features.
  - o Analyze frequencies for categorical features.
  - o Summarize your findings in a detailed report.

1.2 Handling Missing Values and Outliers

- **Objective**: Identify and handle missing data points and outliers using statistical techniques.
- **Instructions**:
  - o Identify missing values in the dataset and propose methods for handling them.
  - o Detect outliers and decide on appropriate methods for treating them.
  - o Justify your chosen methods in a detailed report.

1.3 Data Visualization

- **Objective**: Create compelling data visualizations to showcase relationships between features and the target variable.
- **Instructions**:
  - o Use scatter plots, histograms, or box plots to visualize the data.
  - o Identify trends and patterns through visualizations.
  - o Include the visualizations and insights in your report.

Task 2: Model Selection and Training (LO1, LO3)

2.1 Model Selection

- **Objective**: Choose appropriate machine learning algorithms suitable for predicting binary outcomes (conversion/no conversion).
- **Instructions**:
  - Describe at least two machine learning algorithms (e.g., Logistic Regression, Random Forest).
  - Justify why these algorithms are suitable for this task.
  - Summarize your choices and reasoning in a report.

2.2 Data Splitting

- **Objective**: Split the data into training and testing sets to evaluate and fine-tune the model's performance.
- **Instructions**:
  - Describe the process of splitting the dataset.
  - Ensure the split is stratified to maintain the proportion of the target variable in both sets.
  - Document the details of the data splitting process.

Task 3: Model Interpretation and Evaluation (LO1, LO2, LO3)

3.1 Model Interpretation

- **Objective**: Explain how the chosen model makes predictions and identify the features that most influence the outcome.
- **Instructions**:
  - Perform feature importance analysis using techniques appropriate to your chosen model.
  - Explain the decision-making process of the model.
  - Summarize the key findings and insights in a detailed report.

3.2 Model Evaluation

- **Objective**: Evaluate the model's performance using various metrics and fine-tune hyperparameters to improve the model.
- **Instructions**:
  - Use metrics such as accuracy, precision, recall, and F1-score to assess the model's performance on the validation set.
  - Fine-tune the hyperparameters and document the tuning process.
  - Present the final evaluation results in a detailed report.

Submission Instructions:
- Address each task comprehensively, demonstrating your understanding of the data analysis process and its application to the HashSysTech Insurance scenario.

- You need to use the provided dataset for this assignment and include the code of your Colab implementation link in your assignment.
- Ensure that your report is clear, well-organized, and visually appealing.
- Prepare a document using the BSBI assignment template available in Canvas.
- Use Harvard referencing style for your bibliography.
- Refer to the Essay-Guide available in Canvas for further instructions.
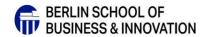- Submit your assignment electronically by the specified deadline.

GRADING DESCRIPTORS: LEVEL 7

EXPERIMENTATION & INNOVATION								
	FAIL			**PASS**				
Threshold Criteria	0-29%	30-39%	40-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Deals with complex issues both systematically and creatively demonstrating self-direction and originality in tackling and solving problems	Little to no ability to use techniques to deal with complex issues systematically (including those of ethics and sustainability) and creatively to solve problems and/or make decisions.	Low utilisation of established techniques to deal with complex issues systematically (including those of ethics and sustainability) and creatively to solve problems and/or make decisions, but with limitations in techniques or approach.	Limited research or advanced scholarship to their area of study by using a range of information and established and advanced techniques	Competent understanding of solving problems, through own research or advanced scholarship displaying a comprehensive understanding of established and advanced techniques	Good understanding of solving problems through own research and advanced scholarship critically selecting and displaying a comprehensive understanding of established and advanced techniques.	Very Good problem-solving skills displaying a comprehensive understanding of techniques applicable to their own research or advanced scholarship	Excellent range of extremely well-developed problem-solving displaying an understanding of techniques applicable to their own research or advanced scholarship beyond which is taught.	Exceptional problem-solving skills with sophisticated evaluation and application of a wide range of advanced information and techniques to undertake projects.
Comprehensive understanding of techniques applicable to their own research or advanced scholarship	Little to no understanding of techniques applicable to their own research or advanced scholarship or their limitations and ambiguities.	Low understanding of techniques applicable to their own research or advanced scholarship including their limitations and ambiguities.	Limited understanding of key techniques applicable to their own research or advanced scholarship including their limitations and ambiguities.	Competent understanding of techniques applicable to their own research or advanced scholarship including their limitations and ambiguities	Good understanding of techniques applicable to their own research or advanced scholarship and a some understanding of more specialised techniques.	Very good understanding of techniques applicable to their own research or advanced scholarship and a some understanding of more specialised techniques.	Excellent understanding of techniques applicable to their own research or advanced scholarship and mastery of some more specialised areas.	Exceptional understanding of techniques applicable to their own research or advanced scholarship and mastery of some more specialised areas.

GRADING DESCRIPTORS: LEVEL 7

RESEARCH & ANALYSIS

Threshold Criteria	FAIL			PASS				
	0-29%	30-39%	40-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Systematic understanding of knowledge, and a critical awareness of current problems and/or new insights, much of which is at, or informed by, the forefront of their academic discipline, field of study or area of professional practice	Little to no knowledge of the subject with limited breadth or depth or deficiencies in major areas or currency.	Low knowledge of the subject lacking coherence, breadth, or detail with only some reference to ideas or arguments at the forefront of any part of the subject.	Limited knowledge to deal with terminology, facts and concepts some of which is informed by the forefront of defined areas of the subject.	Competent knowledge of ideas or arguments at the forefront of any part of the subject sufficient to deal with current issues in the discipline, generally more descriptive than critical or analytical.	Good knowledge of ideas or arguments at the forefront of any part of the subject showing a clear, critical insight into the discipline as whole and current issues/problems.	Very good knowledge of ideas or arguments at the forefront of the subject some of which are significantly beyond what has been taught and show a critical insight into the discipline and current issues/problems.	Excellent knowledge of ideas or arguments at the forefront of the subject many of which are significantly beyond what has been taught and show a critical insight into the discipline and current issues/problems.	Exceptional knowledge of ideas or arguments at the forefront of the subject most of which are significantly beyond what has been taught and show a critical insight into the discipline and current issues/problems.
Conceptual understanding that enables the student to display originality in the application of knowledge	Little to no conceptual understanding or argument and a focus on descriptive explanations which do not comment on arguments of others or alternative views.	Low conceptual understanding and arguments are weak or poorly constructed, and the work does not critically evaluate the arguments of others or consider alternative views.	Limited conceptual understanding and argument construction with critical evaluation of alternative views or comment on advanced scholarship.	Competent conceptual understanding and argument construction with critical evaluation of a range of views and consistent engagement with advanced scholarship.	Good conceptual understanding which critically evaluate and synthesise other views and information with a thoughtful interpretation of advanced scholarship.	Very good conceptual understanding which systematically synthesises a wide range of views with a critical insight into advanced scholarship.	Excellent conceptual understanding which critically apply a wide range of views through a perceptive use of advanced scholarship.	Exceptional conceptual understanding of publishable quality with systematic engagement and usage of advanced scholarship.

GRADING DESCRIPTORS: LEVEL 7

ENGAGING WITH PRACTICE								
	FAIL			**PASS**				
Threshold Criteria	0-29%	30-39%	40-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Practical understanding of how established techniques of research and enquiry are used to create and interpret knowledge in the discipline	Little to no evidence of background investigation, analysis, research, enquiry, ethical awareness, and/or study.	Low evidence of background investigation, analysis, research, enquiry, ethical awareness, and/or study.	Limited background investigation, analysis, research, enquiry , ethical awareness, and/or study using established techniques, with the ability to extract relevant points.	Competent investigation, analysis, research, enquiry, ethical awareness, and/or study using established techniques accurately, and can critically appraise and use academic sources.	Good background investigation, analysis, research, enquiry, ethical awareness, and/or study using established techniques accurately, and possesses a well-developed ability to critically appraise a wide range of sources.	Very good, independent, extensive and appropriate investigation, analysis, research, enquiry, ethical awareness, and/or study beyond the usual range, and critically evaluates this to advance the work and/or direct arguments.	Excellent independent, extensive and appropriate investigation, analysis, research, enquiry, ethical awareness, and/or study well beyond the usual range, and critically evaluates this to advance the work and/or direct arguments.	Exceptional investigation, analysis, research, enquiry, ethical awareness, and/or study which demonstrates carefully considered depth and breadth and critically synthesises this to advance the work and/or direct arguments.
Originality in the application of knowledge	Little to no technical, creative or artistic skills related to their area of study.	Low technical, creative or artistic skills related to their area of study.	Limited technical, creative or artistic skills required for area of study.	Competent technical, creative or artistic skills required for area of study.	Good technical, creative or artistic skills required for area of study.	Very good range of technical, creative or artistic skills.	Excellent range of technical, creative or artistic skills	Exceptional range of technical, creative or artistic skills
Independently advance your own knowledge and understanding, and to develop new skills to a high level.	Little to no contribution to group activity and/or undertaking further training at a high/advanced level.	Low contribution to group activity and/or undertaking further training at a high/advanced level.	Limited contribution to group activity and/or undertaking further training at a high/advanced level.	Competent contribution to group activity and/or independently undertakes further training at a high/advanced level.	Good contribution to group activity and/or independently undertakes further training at a high/advanced level with an understanding of team roles	Very good contribution to group activity and/or independently undertakes further training at a high/advanced level with an understanding of team roles	Excellent contribution to group activity and/or independently undertakes further training at a high/advanced level with teamwork and leadership	Exceptional contribution to group activity and/or independently undertakes further training at a high/advanced level with teamwork and strong leadership.

GRADING DESCRIPTORS: LEVEL 7

REALISATION & COMMUNICATION								
	FAIL			**PASS**				
Threshold Criteria	**0-29%**	**30-39%**	**40-49%**	**50-59%**	**60-69%**	**70-79%**	**80-89%**	**90-100%**
Communicate information, ideas, problems and solutions to both specialist and non-specialist audiences.	Little to no clarity in the communication of ideas, problems and solutions to audiences.	Low clarity in the communication of ideas, problems and solutions to audiences.	Limited clarity in the communication of ideas, problems and solutions to audiences.	Competent communication of ideas, problems and solutions to audiences.	Good, confident and clear communication of ideas, problems and solutions to audiences in a range of means / media.	Very good, confident and clear communication of ideas, problems and solutions to audiences in a range of means / media.	Excellent communication of ideas, problems and solutions to audiences in a range of means / media.	Exceptional communication of ideas, problems and solutions to audiences in a range of means / media.

GRADING DESCRIPTORS: LEVEL 7

PERSONAL & PROFESSIONAL CONNECTIVITY								
	FAIL			**PASS**				
Threshold Criteria	0-29%	30-39%	40-49%	50-59%	60-69%	70-79%	80-89%	90-100%
Independently advance your own knowledge and understanding, and develop new skills to a high level.	Little to no contribution to group activity and/or undertaking further training at a high/advanced level.	Low contribution to group activity and/or undertaking further training at a high/advanced level.	Limited contribution to group activity and/or undertaking further training at a high/advanced level.	Competent contribution to group activity and/or independently undertakes further training at a high/advanced level.	Good contribution to group activity and/or independently undertakes further training at a high/advanced level with an understanding of team roles	Very good contribution to group activity and/or independently undertakes further training at a high/advanced level with an understanding of team roles	Excellent contribution to group activity and/or independently undertakes further training at a high/advanced level with teamwork and leadership	Exceptional contribution to group activity and/or independently undertakes further training at a high/advanced level with teamwork and strong leadership.
Qualities and transferable skills necessary for employment requiring: (a) the exercise of initiative, ethical and personal responsibility (b) decision-making in complex and unpredictable contexts	Little to no ability to manage learning and/or exercise initiative, ethical and personal responsibility and/or decision-making in complex and unpredictable situations	Low ability to manage learning and/or exercise initiative, ethical and personal responsibility and/or decision-making in complex and unpredictable situations	Limited ability to manage learning and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations	Competent ability to manage learning, and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations	Good ability to systematically manage learning, and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations	Very good ability to systematically manage learning, and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations.	Excellent ability to manage learning on own initiative, and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations	Exceptional ability to manage learning on own initiative, and exercise initiative, ethical and personal responsibility, and decision-making in complex and unpredictable situations
	Little to no use of appropriate terminology, limited vocabulary and many errors in spelling, grammar and syntax.	Low use of appropriate terminology, with many errors in spelling, vocabulary and syntax.	Limited expression, style and appropriate vocabulary with errors in spelling, grammar and syntax which affect understanding.	Competent expression, style, and appropriate vocabulary with some errors in spelling, grammar and syntax which do not affect understanding.	Good expression, style and appropriate vocabulary with some errors in spelling, grammar and syntax.	Very good expression, style and appropriate vocabulary with minimal errors in spelling, grammar and syntax.	Excellent expression, style and appropriate vocabulary with minimal errors in spelling, grammar and syntax.	Exceptional expression, style and appropriate vocabulary with no errors in spelling, grammar and syntax.
	Little to no evidence of basic numeracy or digital literacy, hardware and software skills	Low evidence of basic numeracy or digital literacy, hardware and software skills competency.	Limited evidence of numeracy or digital literacy, hardware and software skills competency.	Adequate evidence of numeracy or digital literacy, hardware and software skills competency.	Good evidence of numeracy or digital literacy, hardware and software skills competency.	Very good evidence of numeracy or digital literacy, hardware and software skills	Excellent evidence of numeracy or digital literacy, hardware and software skills competency.	Exceptional evidence of numeracy or digital literacy, hardware and software skills competency.

					competency.		
	Does not demonstrate achievement of professional competence when assessed against the requirements of a professional, statutory or regulatory body (PSRB).			The student has demonstrated achievement of professional competence when assessed against the requirements of a PSRB.			
	Inaccurate use of terminology with limited vocabulary and many errors in spelling, grammar and syntax. Inaccurate terminology, with many errors in spelling, vocabulary and syntax.			The student has adhered to the appropriate rules and/or conventions set by regulators or the industry.			