<table>
<tr><td>Name of the program:</td><td colspan="3">MSc program in Applied Statistics</td></tr>
<tr><td>Department:</td><td colspan="3">Department of Statistics</td></tr>
<tr><td>Semester/Year/Intake</td><td colspan="3">Semester 2/Year2/2022</td></tr>
<tr><td>Year:</td><td colspan="3">2024</td></tr>
<tr><td>Course Code:</td><td colspan="3">STA 529 2.0</td></tr>
<tr><td>Course Name:</td><td colspan="3">Data Mining</td></tr>
<tr><td>Credit Value:</td><td colspan="3">2.0</td></tr>
<tr><td>Core/Optional</td><td colspan="3">Core</td></tr>
<tr><td rowspan="2">Hourly Breakdown</td><td>Theory</td><td>Practical</td><td>Independent Learning</td></tr>
<tr><td>20</td><td>10</td><td>70</td></tr>
<tr><td colspan="4">Course Aim/Intended Learning Outcomes:
At the completion of this course student will be able to

➢ Apply suitable classification/regression techniques to solve a given problem
➢ Validate the model fitted using a suitable method
➢ Extract important association rules from a given dataset
➢ Extract important rules across time/position from a given dataset
➢ Extract patterns from a large unstructured text data</td></tr>
<tr><td colspan="4">Course Content: (Main topics, Sub topics)

➢ Introduction to Data Mining
➢ Classification and Regression Methods in Data Mining
K-Nearest Neighbour, Linear and Quadratic Discriminant Analysis, Decision Trees, Bagging, Random Forest, Boosting, Support Vector Machine
➢ Model validation
K-fold Cross validation, ROC Analysis, Contingency Table Based Measures
➢ Association Rule Mining
➢ Introduction to Text Mining</td></tr>
<tr><td colspan="4">Teaching /Learning Methods: Lectures and student-centered teaching learning methods
Mode of Delivery:
All lectures will be delivered using online teaching methods till the university grants permission to conduct face-to-face lectures for postgraduate students.</td></tr>
<tr><td colspan="4">Assessment Strategy:</td></tr>
<tr><td colspan="2">Continuous Assessment</td><td colspan="2">Final Assessment</td></tr>
</table>

<table>
<tr><td colspan="2">30%</td><td colspan="3">70%</td></tr>
<tr><td colspan="2">Details: quizzes %, mid-term %, other % (specify)

0 % 100% 0%</td><td>Theory (%)

50%</td><td>Practical (%)

50%</td><td>Other (%)(specify)

0</td></tr>
<tr><td colspan="5">References/Reading Materials:

➢ Data Mining and Analysis: Fundamental Concepts and Algorithms, Mohommed J. Zaki, Wagner Meira Jr. Cambridge University Press New York, NY, USA
➢ Data Mining Concepts and Techniques 3rd Edition, Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann Publishers 225Wyman Street,Waltham, MA 02451, USA
➢ R and Data Mining: Examples and Case Studies Yanchang Zhao http://www.RDataMining.com</td></tr>
</table>