

# CONTENTS

---

<b>1. INTRODUCTION .....</b>	<b>1</b>
<b>2. PROBLEM STATEMENT .....</b>	<b>3</b>
2.1 STAKEHOLDERS' QUESTIONS .....	3
2.2 QUESTIONS REFORMULATION.....	4
2.2.1 <i>Learning Profiles</i> .....	4
2.2.2 <i>Dropout prediction</i> .....	4
2.3 THE DATA SCIENCE SOLUTION .....	4
<b>3. PLANNING, ANALYSIS AND DESIGN .....</b>	<b>5</b>
3.1 PLAN DESIGN .....	5
3.2 PLAN FOLLOW-UP .....	5
<b>4. BACKGROUND RESEARCH .....</b>	<b>7</b>
4.1 STATISTICAL AND VISUALIZATION APPROACHES.....	7
4.2 DATA MODELING APPROACHES .....	7
<b>5. IMPLEMENTATION.....</b>	<b>9</b>
5.1 DATA ASSESSMENT .....	9
5.2 METHODOLOGY AND TOOLS.....	10
5.3 DATA BATCH PROCESSING .....	11
5.4 FEATURE ENGINEERING .....	11
5.5 MODELS TO UNDERSTAND BEHAVIORS.....	12
5.6 PREDICTION MODEL .....	12
<b>6. RESULTS AND DISCUSSION .....</b>	<b>15</b>
6.1 EVALUATION CRITERIA .....	15
6.2 MODELS TO UNDERSTAND BEHAVIORS.....	15
6.3 PREDICTION MODEL .....	16
6.4 DISCUSSION .....	16
<b>7. CONCLUSIONS AND FUTURE WORK .....</b>	<b>17</b>
7.1 FUTURE WORK .....	17
<b>REFERENCES.....</b>	<b>19</b>
<b>8. APPENDIX A – PROJECT PLAN AND FOLLOW-UP .....</b>	<b>21</b>
8.1 GANTT CHART .....	21
8.2 TASKS DESCRIPTION .....	21
<b>9. APPENDIX B – COURSE FILES STRUCTURE .....</b>	<b>23</b>