

# Table of Content

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Research questions . . . . .	2
1.3	Objectives . . . . .	2
1.3.1	General objective . . . . .	2
1.3.2	Specific objectives . . . . .	2
<b>2</b>	<b>Background and Related work</b>	<b>3</b>
2.1	Twitter and elections . . . . .	3
2.1.1	Chile . . . . .	5
2.2	Measuring influence . . . . .	5
2.2.1	Volumetry . . . . .	6
2.2.2	Centrality in Social Network Analysis . . . . .	6
<b>3</b>	<b>Methods</b>	<b>7</b>
3.1	Data collection . . . . .	7
3.1.1	Candidates data . . . . .	7
3.1.2	Twitter data . . . . .	8
3.2	Feature engineering . . . . .	9
3.2.1	Twitter metrics . . . . .	10
3.2.2	Sentiment analysis . . . . .	12
3.2.3	Network analysis . . . . .	13

3.2.4	Custom political PageRank . . . . .	14
3.3	Validation . . . . .	17
3.3.1	Preprocessing . . . . .	17
3.3.2	Base features . . . . .	19
3.3.3	Correlation . . . . .	20
3.3.4	Regression . . . . .	21
3.3.5	Classification . . . . .	23
3.3.6	Political PageRank . . . . .	28
<b>4</b>	<b>Results</b>	<b>29</b>
4.1	RQ1: Twitter influence features assessment . . . . .	29
4.1.1	Descriptive analysis . . . . .	29
4.1.2	Correlation . . . . .	32
4.1.3	Regression . . . . .	34
4.1.4	Classification . . . . .	38
4.2	RQ2: Modified PageRank . . . . .	45
<b>5</b>	<b>Conclusions</b>	<b>52</b>
5.1	Description, Explanation or Prediction? . . . . .	53
	<b>Bibliography</b>	<b>55</b>
	<b>Annexed</b>	<b>59</b>