

# Table of Content

<b>1. Introduction</b>	<b>1</b>
1.1. Time Series . . . . .	1
1.1.1. Time Series in general . . . . .	1
1.1.2. Time Series in Astronomy . . . . .	2
1.2. Problem statement . . . . .	4
1.3. Hypothesis . . . . .	4
1.4. Objectives . . . . .	4
General Objective . . . . .	5
Specific Goals . . . . .	5
1.5. Methodology . . . . .	5
1.6. Contributions . . . . .	6
1.7. Outline . . . . .	7
<b>2. Related work</b>	<b>8</b>
2.1. Key concepts on Time Series representation . . . . .	8
2.1.1. Time Series data . . . . .	8
2.1.2. Basis of representation method . . . . .	9
2.1.3. Data mining applications. . . . .	9
2.2. Similarity measures . . . . .	10
2.2.1. Minkowski distance . . . . .	11
2.2.2. Cosine Similarity . . . . .	11
2.2.3. Dynamic Time Warping (DTW) distance . . . . .	12
2.2.4. Edit Distance . . . . .	13
2.2.5. Similarity measures for irregular-multivariate Time Series . . . . .	14
2.3. Efficient similarity search . . . . .	14
2.3.1. lower bounding distance . . . . .	15
2.3.2. indexing strategies . . . . .	16
2.3.2.1. One-dimensional data . . . . .	16
2.3.2.2. High-dimensional data . . . . .	16
2.3.2.3. Irregular sampled and multivariate data . . . . .	16
2.4. Time Series Representations . . . . .	17
2.4.1. Regular-Univariate data approaches . . . . .	17
2.4.1.1. Feature extraction . . . . .	17
2.4.1.2. Piecewise approximation . . . . .	18
2.4.1.3. Symbolic representation and Bag-of-Words . . . . .	21
2.4.2. Regular-Multivariate data approaches . . . . .	22
2.4.3. Irregular-Univariate data approaches . . . . .	23

2.4.4.	Irregular-Multivariate data approaches . . . . .	25
2.5.	Discussion . . . . .	27
2.5.1.	Data mining application . . . . .	27
2.5.2.	Research directions . . . . .	29
<b>3.</b>	<b>Information Retrieval theory for Time Series</b>	<b>30</b>
3.1.	Time Series Transformation . . . . .	30
3.1.1.	Document representation . . . . .	30
3.1.2.	Numerosity reduction . . . . .	32
3.2.	Vector space model . . . . .	32
3.2.1.	Bag-of-Words model . . . . .	33
3.2.2.	TF-IDF model . . . . .	34
3.2.3.	Sub-linear TF weights . . . . .	35
3.2.4.	Cosine normalization . . . . .	35
3.2.5.	Class based TF-IDF . . . . .	35
3.3.	Compact vectors . . . . .	36
3.3.1.	Latent Semantic Analysis (LSA) . . . . .	36
3.3.2.	One-way Analysis of Variance (ANOVA) . . . . .	38
3.3.3.	One-way Multivariate ANOVA (MANOVA) . . . . .	40
3.4.	Summary . . . . .	42
<b>4.</b>	<b>A new representation method for astronomical Time Series</b>	<b>43</b>
4.1.	Irregular time Series . . . . .	43
4.1.1.	Pattern extraction . . . . .	43
4.1.2.	Pattern segmentation . . . . .	45
4.2.	Irregular Bag-of-Pattern Feature approach . . . . .	45
4.3.	Generalized approach for Multivariate cases . . . . .	47
4.3.1.	Compact representation using MANOVA . . . . .	48
4.3.2.	Compact representation using LSA . . . . .	49
4.4.	Extensions on Irregular Bag-of-Pattern Feature . . . . .	50
4.4.1.	Including Multi-Quantity representation . . . . .	50
4.4.2.	Including Multi-Resolution representation . . . . .	54
4.5.	Discussion . . . . .	54
4.5.1.	Proposed methods . . . . .	55
4.5.2.	Computational complexity . . . . .	56
4.5.3.	Data Mining application . . . . .	56
<b>5.</b>	<b>Experimental Evaluation</b>	<b>57</b>
5.1.	Implementation . . . . .	57
5.2.	Dataset . . . . .	58
5.3.	Literature methods for comparison . . . . .	61
5.4.	Parameter explorations . . . . .	61
5.4.1.	Multi-Quantity search . . . . .	63
5.4.2.	Optimal Levels of resolution . . . . .	65
5.5.	Classification Comparison with the State-of-the-art . . . . .	66
5.5.1.	Optimal compact technique . . . . .	66
5.5.2.	Classification on full test set . . . . .	68

5.5.3. Extra case study on classification . . . . .	70
5.6. Similarity search experiments . . . . .	73
5.7. Computing time experiments . . . . .	73
<b>6. Conclusion</b>	<b>75</b>
6.1. Contributions . . . . .	75
6.2. Discussion . . . . .	76
6.2.1. Proposed approach . . . . .	76
6.2.2. Representation method performance . . . . .	77
6.2.3. Classification comparissons . . . . .	78
6.2.4. Similarity search . . . . .	80
6.2.5. Computational complexity . . . . .	80
6.3. Future work . . . . .	81
<b>Bibliography</b>	<b>84</b>
<b>Annexed</b>	<b>93</b>
Additional classification figures . . . . .	93