Universidade do Minho
Escola de Engenharia
Departamento de Sistemas de Informação

Bruno Rafael Azevedo Ferreira

*Emotions recognition based on sensor fusion and machine learning techniques*

February 2018
Emotions recognition based on sensor fusion and machine learning techniques
ABSTRACT

Technological evolution and the resulting increase in its importance and dependence for the organizations and society have led, in recent years, to the significant growth in the volume, variety and speed of generated data. In this sense, the different electronic devices, sensors, people and infrastructures increasingly communicate with each other, exchanging and generating enormous amounts of data, with multiple levels of complexity, at different speeds.

Nowadays, the data generated by these devices is of great importance and can be used in many different areas and for a wide range of purposes. This is where the concepts of Big Data, Sensor Fusion and Machine Learning gain importance. Big Data allows for efficient processing and storage of data that is generated daily, Sensor Fusion (Data Fusion in sensor networks) ensures, through its algorithms, processes and methods, that a better representation of the real world is obtained, thus improving a particular context to be analyzed, something that could not be achieved by only one type of data, and finally, Machine Learning allows the machines to acquire knowledge through artificial intelligence.

These three concepts are very important and relevant to the context of this dissertation, which consists of the emotional identification of a person in the context of the automobile industry, since the psychological state of a human being has a great impact on his behaviour.

Thus, this dissertation initially seeks the development of a Sensor Fusion platform that allows the collection of synchronized data of audio and video, followed by the creation of a machine learning project to classify the emotional states of the people through the collected data and, finally, the fusion of the models developed to guarantee a better representation of the emotional state.

In this document, it is performed an identification of the objectives, motivations and scope of the project, followed by a literature review of the topics to be addressed in the dissertation (with the objective of understanding and having a more concrete and clear vision of these themes). After that, it is done a study of the methodological approaches that are going to be used in the project, and finally, the planning of the entire detailed project with all the tasks that are to be carried out is developed.

**Keywords:** Sensor Fusion, Big Data, Machine Learning, Emotions Recognition
RESUMO

A evolução tecnológica e o resultante aumento da sua importância e dependência para as organizações e sociedade levaram, nos últimos anos, ao crescimento significativo do volume, variedade e velocidade de dados gerados. Neste sentido, cada vez mais os diferentes dispositivos eletrônicos, sensores, pessoas e infraestruturas comunicam entre si, trocando e gerando enormes quantidades de dados, com múltiplos níveis de complexidade, a diferentes velocidades.

Os dados gerados por estes dispositivos apresentam grande importância na sociedade atual, podendo ser utilizados nas mais diversas áreas e para os mais diversos fins. É aqui que os conceitos de Big Data, Sensor Fusion e Machine Learning ganham importância. O Big Data permite um eficiente processamento e armazenamento dos dados que são gerados diariamente, o Sensor Fusion (Data Fusion em redes de sensores) garante, através dos seus algoritmos, processos e métodos, que se obtenha uma melhor representação do mundo real, melhorando assim um determinado contexto a ser analisado, algo que não poderia ser alcançado por apenas um tipo de dados e, por fim, o Machine Learning permite que as máquinas adquiram conhecimento através de inteligência artificial.

Estes três conceitos são bastante importantes e relevantes para o contexto desta dissertação, esta que consiste na identificação emocional de uma pessoa no contexto na indústria automóvel, visto que o estado psicológico de um ser humano apresenta grande impacto na sua condução.

Assim, esta dissertação procura, inicialmente, o desenvolvimento de uma plataforma de Sensor Fusion que permita recolher dados sincronizados de áudio e vídeo, seguido da criação de um projeto de machine learning para se classificar os estados emocionais das pessoas através dos dados recolhidos e, por fim, a fusão dos modelos desenvolvidos para se garantir uma melhor representação do estado emocional.

Neste documento é feita uma identificação dos objetivos, motivações e enquadramento do projeto, seguido da elaboração de uma revisão de literatura sobre os temas a serem abordados na dissertação (com o objetivo de se compreender e ter uma visão mais concreta e clara destes temas). Após isto, faz-se um estudo das abordagens metodológicas a serem utilizadas no projeto e, para terminar, é desenvolvido o planeamento de todo o projeto detalhado com todas as tarefas a serem realizadas.

**Palavras-Chave:** Sensor Fusion, Big Data, Machine Learning, Emotions Recognition
## CONTENTS

1 INTRODUCTION
   1.1 Scope and Motivation 1
   1.2 Project Objectives 1
   1.3 Document Outline 2

2 STATE OF THE ART
   2.1 Bibliographic Search Strategy 3
   2.2 Sensor Fusion 5
      2.2.1 Definition 5
      2.2.2 Sensor Fusion Architectures 9
      2.2.3 Types of Sensors 12
      2.2.4 Algorithms 13
   2.3 Big Data 22
      2.3.1 Definition 22
      2.3.2 Types of NoSQL Databases 25
   2.4 Machine Learning 29
      2.4.1 Definition 29
      2.4.2 Classification 31
      2.4.3 Neural Networks 33
   2.5 SOTA Relevant Sections 35

3 PROBLEM STATEMENT 39

4 WORKPLAN 40

REFERENCES 44
| Figure 1 | Research Method | 4 |
| Figure 2 | Keywords used on the research process | 5 |
| Figure 3 | Centralized Fusion Architecture (adapted from (C. Li, Azzam, & Fernández-Steeger, 2016)) | 10 |
| Figure 4 | Distributed Fusion Architecture (adapted from (C. Li et al., 2016)) | 11 |
| Figure 5 | Groups of Kalman filter equations (adapted from (Baskaran & Gangadharan, 2016; Souza, 2017)) | 16 |
| Figure 6 | Sequential Monte Carlo methods steps (created based on (Baskaran & Gangadharan, 2016)) | 18 |
| Figure 7 | Differences between fuzzy and crisp sets | 20 |
| Figure 8 | V's based on different authors (created based on (d. Almeida & Bernardino, 2015; Diebold, 2012; Fan & Bifet, 2013; Luna Dong & Srivastava, 2013; Sagiroglu & Sinanc, 2013; Sharma, 2015; u. d. Khan, Uddin, & Gupta, 2014)) | 23 |
| Figure 9 | Additional Big Data Characteristics (created based on (Krishnan, 2013a)) | 24 |
| Figure 10 | CAP Theorem (created based on (J. Han, E, Le, & Du, 2011)) | 26 |
| Figure 11 | Bias-Variance dilemma (adapted from (Fortmann-Roe, 2012)) | 31 |
| Figure 12 | Machine Learning Process of this dissertation | 32 |
| Figure 13 | Artificial Neural Network with one hidden layer | 34 |
| Figure 14 | Different types of Neural Networks (created based on (Cortez, 2010; Haykin, Haykin, Haykin, & Haykin, 2009; Rocha, Cortez, & Neves, 2007; Schmidhuber, 2015)) | 35 |
| Figure 15 | Gantt Diagram with all the tasks | 43 |
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Relevant Definitions by Chronological Order</td>
<td>6</td>
</tr>
<tr>
<td>Table 2</td>
<td>Relevant Databases Analysis</td>
<td>28</td>
</tr>
<tr>
<td>Table 3</td>
<td>Relevant Literature</td>
<td>36</td>
</tr>
<tr>
<td>Table 4</td>
<td>Schedule</td>
<td>40</td>
</tr>
<tr>
<td>Table 5</td>
<td>Dissertation's Work Plan</td>
<td>41</td>
</tr>
</tbody>
</table>
INTRODUCTION

1.1 SCOPE AND MOTIVATION

Human emotions represent an external manifestation of a person's inner emotional state. Thus, it represents a powerful communication signal that can be transmitted by various sources of information, such as facial, vocal and corporal gestures (Mauss & Robinson, 2009).

With the development of artificial intelligence, machines are increasingly becoming able to learn through models developed through Machine Learning techniques. The Machine Learning technologies are very present in modern society, and can be used for the most diverse purposes.

Machine Learning seeks to create algorithms that can be applied to an existing set of data to improve and refine its capabilities to deal with future data, by helping machines discover patterns of information that correspond to a specific result. (Stanton, 2013).

On the other hand, Sensor Fusion refers to the process of successfully combining data from several homogeneous or heterogeneous sensors (Geng, Liang, Liu, & Alsaadi, 2018) thus obtaining a better representation of the real world.

The project is being developed as part of a curricular internship at Bosch Car Multimedia Portugal S.A. and seeks to take advantage of the techniques of Machine Learning and Sensor Fusion to classify the human emotions of people in a vehicle due to behavioral changes that can be caused by negative emotions, such as stress and violence.

Although this is the main focus of the dissertation, since there were two case studies in the beginning, both are discussed in this document.

1.2 PROJECT OBJECTIVES

An efficient goal definition is critical to the development of any project. In this sense, and like all projects, this dissertation presents well defined objectives and results. The main objective of this dissertation is to identify violence and stress of the different people who are in a car.

In a first phase, it will be collected information about the different concepts that are relevant to this dissertation, such as relevant information on sensor fusion and machine learning. Then, a sensor fusion system that is capable of collecting data from cameras and microphones in a synchronized way
is expected to be developed. Subsequently, the same data should be used in the context of facial and voice recognition, followed by the classification of several emotions. The use of different machine learning algorithms for each of the data types used is expected.

Finally, the two generated models that classify the different emotions through audio and video will be fused to ensure a better representation of the emotional state of a particular person.

1.3 DOCUMENT OUTLINE

This document is divided into 4 chapters which allows a better understanding of the topic to be addressed in this dissertation. Thus:

- **Chapter 1** presents the Introduction of the topic to be developed, where the Scope and Motivation, the Objectives and Expected Outcomes and, finally, the structure of the report of this dissertation are identified;

- **Chapter 2** presents all the State of the Art developed for this dissertation. Thus, this chapter starts by identifying the methodological approaches and the research method used for the State of the Art development, followed by the three main concepts relevant for this dissertation (Sensor Fusion, Big Data and Machine Learning). The Sensor Fusion point is divided into four sub-chapters: the Definition of Sensor Fusion, its architectures, the Types of Sensors used in this dissertation, and finally, different algorithms used in this context. The Big Data chapter point presents two sub-chapters: the Definition of Big Data and the different types of NoSQL databases. Finally, the Machine Learning point has 3 sub-chapters: the Definition of Machine Learning, where is described what is Machine Learning, the different types of Machine Learning and, at last, the problems related to it. The second point consists on a general analysis of the Machine Learning Classification method, where is described what is classification, its challenges and its steps. Finally, the last point talks about Neural Networks, for example, what is a Neural Network, its characteristics and the two learning procedures of the Neural Networks.

- **Chapter 3** presents the Problem Statement where the main problem of this dissertation is identified, with a detailed technical explanation;

- **Chapter 4** consists of the Work Plan, that is, in this chapter an approach is presented to the project management. This approach presents a detailed description of the planning that will be elaborated, with all the project activities and estimated deadlines for each activity.
2

STATE OF THE ART

2.1 BIBLIOGRAPHIC SEARCH STRATEGY

To develop an efficient state of the art that allows the reader to analyze and understand the content present in it, it is necessary to follow a well-defined and structured research and documentation process, which must be able to support the developed work.

In this sense, for this document, it was necessary to use several search engines and informational databases, that encompassed several published scientific articles, chapters of books, dissertations, conference proceedings, scientific journals, among others. It was also taken into account the diverse documentation provided by the advisor of this dissertation.

The websites most used for the development of this document were:

- Science Direct
- Google Scholar
- Google
- Springer
- Scopus
- Research Gate
- RepositoriUM

In order to use these websites, it was necessary to choose and develop, at an early stage, the critical concepts related to this dissertation.

Since a lot of documentation was collected, and not all of them would be relevant to the inserted context, a method of filtering it was established. This method consisted of classifying the documentation as very relevant, not very relevant and irrelevant.

- **Very Relevant**: Careful and thorough reading of the documentation, making use of a sense of critical analysis;

- **Not very relevant**: Reading and superficial analysis of the abstract and, if necessary, of the conclusions of the document;

- **Irrelevant**: Discard documentation because it is not useful for the topic of the dissertation.
The methods used to classify the different collected documents can be found in figure 1.

From the observation of figure 1, we can conclude that if a given document was made available by the advisor, the abstract would be read to verify if it was relevant to the dissertation. If this was not relevant, it would be classified as irrelevant literature and, if relevant, the introduction and conclusion would be read. In this sense, if the introduction and conclusion were shown to be very relevant, this documentation would be classified as very relevant and, if both were not very relevant, the document would be classified as relevant or not very relevant.

In the case of an autonomous search, the documentation was searched in the different search engines presented above. Thus, initially, the keywords related to the theme of this dissertation were identified, followed by the identification of the type of documentation (dissertation, articles, among others), to verify if it did not contain false or not scientifically accepted information. If it was, the year of
publication was analyzed to see if the concepts presented in it were not outdated and, if it was recent, the abstract would be read, followed by the same classification presented when the documents were made available by the advisor and if the document was not recent, the quantity of citations would be verified to analyze if it was classified as irrelevant or if it had relevant information to be approached, that is, if it is a document very old and without any type of citations, we can conclude that it does not have information that is relevant to the topic to be studied, but if it presented many citations, the information made available by it may be crucial for the project.

In this sense, we can overcome the problem of a document being old and not having been analyzed, since some may present recent and relevant information. If it was not a scientific document such as a dissertation, article, conference proceedings, among others, it would be verified whether it was a book or the chapter of a book, if it were not, the document would be immediately discarded because it could contain fallacious information. If it was, the chapter summary would be read and, depending on the level of relevance, it would be classified as very relevant or insignificant.

Finally some keywords were used to make an efficient research process. These keywords can be analyzed on Figure 2.

![Figure 2: Keywords used on the research process](image)

2.2  SENSOR FUSION

2.2.1  Definition

Over the last few years, there has been a great deal of interest in developing systems capable of using multiple sources of sensorial information. Thus, the concepts and the terminology related to data fusion systems (Sensor Fusion, Data Fusion, among others) are a fundamental research topic that has received, during the recent years, significant attention in the literature (Gravina, Alinia, Ghasemzadeh,
There are various kinds of sensor and information fusion systems. However, all of them have a common characteristic, which consists in the information originated from a number of homogeneous or heterogeneous sources being joined to reduce the quantity and dimensionality of data, and to obtain information with higher quality. (D. Hall & Llinas, 2001a).

Nowadays, these systems are widely used in many fields such as robotics, space, bio-medical, transportation, economics and financial information systems, sensor networks, video and image processing, intelligent system design, among others (Khaleghi, Khamis, Karray, & Razavi, 2013; Tan, Shen, Liu, Al- saedi, & Ahmad, 2017). In this way, a lot of concepts related to the information fusion systems (for example: Data Fusion, Sensor Fusion, Information Fusion, among others) have been extensively used as a reference to a diversity of technologies, techniques, systems, and applications that collect, combine and use data derived from diverse information sources (Rothman & Denton, 1991).

As mentioned above, distinct authors present different perspectives and definitions of concepts related to the data fusion systems area, since this topic covers an enormous variety of topics and technologies. In this sense, there have been several attempts to define and categorize fusion terms and techniques.

In order to clarify what was previously mentioned, different definitions for Information Fusion, Data Fusion and Sensor Fusion/Multisensor Fusion will be discussed in this chapter (Table 1). The descriptions are presented in a chronological way, so we can analyze how these concepts have evolved over a period of time, through different perspectives and opinions.

However, all the definitions for this concepts have the same, or similar, purpose and goal, which are to achieve improved accuracy and more grounded inferences than could be obtained by the use of a single sensor (better qualitatively and quantitatively decisions and/or actions), help to understand a particular scenario (Abidi & Gonzalez, 1992), maximize the useful information content (Starr & Desforges, 1998), refined estimates of parameters, characteristics, events, and behaviors for observed entities in an observed field of view (D. Hall & Llinas, 2001b), effective support for human or automated decision making (Boström et al., 2017), among others.

In addition, and as we will verify, some authors present very similar descriptions, which led to the rejection of some of them.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Year</th>
<th>Authors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Fusion</td>
<td>1987</td>
<td>JDL</td>
<td>Process related to the association, correlation and combination of data and information from one or more sources (White, 1991).</td>
</tr>
</tbody>
</table>

(Table 1 continued on next page)
<table>
<thead>
<tr>
<th>Concept</th>
<th>Year</th>
<th>Authors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Fusion</td>
<td>1988</td>
<td>Richardson and Marsh</td>
<td>Process through which data from a multitude of sensors is used to create an optimal estimate of a specified state vector (Richardson &amp; Marsh, 1988).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1990</td>
<td>Waltz and Llinas</td>
<td>Combination of elements of raw data from various sources into one set of significant information with a greater benefit than the sum of the contributing parts. In a technological perspective, it involves the integration and application of several traditional disciplines and new areas of engineering to achieve the fusion of data (Waltz &amp; Llinas, 1990).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1992</td>
<td>Abidi and Gonzalez</td>
<td>It deals with the synergistic combination of information obtained from different knowledge sources (sensors, among others) (Abidi &amp; Gonzalez, 1992).</td>
</tr>
<tr>
<td>Sensor Fusion</td>
<td>1992</td>
<td>Hall</td>
<td>Seeks to combine data collected from several sensors, in order to make inferences that could not be performed by a single sensor (D. L. Hall, 1992).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1994</td>
<td>DSTO</td>
<td>It consists in a multilevel process that deals with the automatic detection, association, correlation, estimation, and combination of information and data derived from one or more sources (DSTO, 1994).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1997</td>
<td>Hall and Llinas</td>
<td>Techniques that combine data from multiple sensors, and related database information, to increase the accuracy and specific inferences that could not be achieved by a single sensor (D. L. Hall &amp; Llinas, 1997).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1997</td>
<td>Goodman, Mahler and Nguyen</td>
<td>It consists in locating and identifying several unknown objects of different types, based on different types of evidence. These evidence is continuously collected by several allocatable sensors having varying capabilities (Goodman, Mahler, &amp; Nguyen, 1997).</td>
</tr>
</tbody>
</table>

(Table 1 continued on next page)
<table>
<thead>
<tr>
<th>Concept</th>
<th>Year</th>
<th>Authors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Fusion</td>
<td>1998</td>
<td>Starr and Desforges</td>
<td>Process that combines data and knowledge from diverse sources to maximize the content of useful information, to improve the reliability or discriminant capability, while minimizing the quantity of data ultimately retained (Starr &amp; Desforges, 1998).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1999</td>
<td>Wald</td>
<td>Formal framework that defines the means and tools for the alliance of data from multiple sources (Wald, 1999).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>1999</td>
<td>Steinberg, Bowman and White</td>
<td>Process that consists in combining data to refine state estimates and predictions (Steinberg, L. Bowman, &amp; White, 1999).</td>
</tr>
<tr>
<td>Information</td>
<td>2001</td>
<td>Llinas</td>
<td>Information process that deals with the association, correlation, and combination of data and information from one or more sensors or sources (D. Hall &amp; Llinas, 2001b).</td>
</tr>
<tr>
<td>Information</td>
<td>2001</td>
<td>Dasarathy</td>
<td>It covers the theory, techniques and tools created and employed for exploring the synergy in the information acquired from various sources (Dasarathy, 2001).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>2001</td>
<td>McGirr</td>
<td>Process that gathers huge amounts of different information into a more comprehensive and easily manageable form (McGirr, 2001).</td>
</tr>
<tr>
<td>Multi-Sensor Data Fusion</td>
<td>2005</td>
<td>Challa, Gulrez, Chaczko and Paranesha</td>
<td>Core element of all networked sensing systems that is used to join/combine complementary information produced by sensors to acquire a more complete picture, or to reduce/manage uncertainty by using sensor information from several sources (Challa, Gulrez, Chaczko, &amp; Paranesha, 2005).</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>2007</td>
<td>Mastrogiovanni</td>
<td>This process consists in maximizing the useful information content collected from heterogeneous sources (Mastrogiovanni, Sgorbissa, &amp; Zaccaria, 2007).</td>
</tr>
</tbody>
</table>
2.2. Sensor Fusion

(Table 1 continued from previous page)

<table>
<thead>
<tr>
<th>Concept</th>
<th>Year</th>
<th>Authors</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Fusion</td>
<td>2007</td>
<td>Boström, Andler, Brohede, Johansson, Karlsson, Laere, Niklasson, Nilsson, Persson and Ziemke</td>
<td>Study of efficient methods for automatically or semi-automatically convert information from multiple sources and different points in time (Boström et al., 2017).</td>
</tr>
<tr>
<td>Sensor Fusion/Multi-Sensor Fusion</td>
<td>2011</td>
<td>Khaleghi, Khamis, Karray and Razavi</td>
<td>Technology that enables the combination of information derived from multiple sources to form a unified picture (Khaleghi et al., 2013).</td>
</tr>
<tr>
<td>Sensor Fusion</td>
<td>2016</td>
<td>Saboonchia, Ozevin and Kabir</td>
<td>Data integration of several sensors to improve the decision-making performance (Saboonchi, Ozevin, &amp; Kabir, 2016).</td>
</tr>
<tr>
<td>Multi-Sensor Data Fusion</td>
<td>2017</td>
<td>Geng, Liang, Liu and Alsaadi</td>
<td>It refers to the process of successfully combining data from several homogeneous or heterogeneous sensors (Geng et al., 2018).</td>
</tr>
</tbody>
</table>

It is also important to analyze that as descriptions advance, new specifications and characteristics are added to the concepts and terminology to the fusion systems, which shows how they have evolved over the previous period.

Having completed the previous exhaustive analysis, and in the business context to which this pre-dissertation is inserted, we can draw our own conclusions about this topic. Thus, I’ve created my own definition for Sensor Fusion. In this sense, Sensor Fusion can be described as “the fusion of many types of data with several levels of complexity, collected at different velocities from a single or multiple sources (sensors) with the main purpose of improving the quality of that data (reducing uncertainty), the decision-making process, and the characterization of a given business context. This process can be performed through different procedures, methods and algorithms chosen according to the data and context inserted. It can also be continuously refined over time, improving its components and characteristics”.

2.2.2 Sensor Fusion Architectures

Like all the fields of study, Sensor Fusion presents two methods that are significant for processing data gathered from several sources (C. Z. Han, Zhu, Duan, et al., 2006). These techniques, that consist
of the combination of information from different sensors, are intended to solve the problem of signal estimation (García-Ligero, Hermoso-Carazo, & Linares-Pérez, 2015).

The first one is called centralized fusion (Figure 3), which consists in transferring all of the measurement data of the individual sensors to a fusion center, to be processed, that is, there is an entity on the fusion center that is in charge for the correlation and fusion of that data. When no faults are displayed by any of the sensors, this method is ideal with certain performance criteria. Conversely, it has poor robustness and reliability when the sensors have faults and it requires costly communication and computing resources to work properly (Sun et al., 2017). Besides that, in this type of fusion, the measurements collected from the sensors are sent to a single node whose main goal is to fuse those measurements, which makes this method very vulnerable to catastrophic failures. (Khosla, Guillochon, & Choe, 2017)

An example of the use of centralized fusion can be verified in the article “Non-parametric Laser and Video Data Fusion: Application to Pedestrian Detection in Urban Environment” in which its authors developed a model based on this method. In this example, the data is collected from two types of sensors (Cameras and Laserscanners) and are merged (associated and tracked) and filtered in the same central block, resulting in the identification of people in the urban environment (Gidel, Blanc, Chateau, Checchin, & Trassoudaine, 2009).

![Figure 3: Centralized Fusion Architecture (adapted from (C. Li et al., 2016))](image)

The second method is called distributed fusion (Figure 4), through which individual sensors provide local state estimates through their measurement data transmitting them, afterwards, to the fusion center to develop a globally optimal or suboptimal state estimate according to certain fusion criteria, which means that the user has the full responsibility for the fusion of that data. This method has parallel
components that facilitate the fault detections and isolation of sensors, becoming, in this sense, a robust, flexible and reliable method. Conversely, it is generally worse when compared with centralized fusion (Sun et al., 2017).

![Distributed Fusion Architecture](adapted_from_C.Li_et_al._2016)

Figure 4: Distributed Fusion Architecture (adapted from (C. Li et al., 2016))

Given the advantages and disadvantages presented for each one of the techniques, we can indicate that the distributed fusion is more efficient in the performance and operation of a system, and it is more robust than the centralized fusion. Thus, both methods are currently very used, however, distributed fusion tends to be preferred because of its computational advantages and practical applications, such as communication data compression, reduction of computational requirements, high fault tolerance, among others (Jin, Lin, & Sun, 2005).

It is also relevant to indicate that both methods are important for Multi-Sensors Systems. These systems allow the use of all the information collected from all the sensors, with maximum performance, and overcome the defects of using only one sensor (N. Li, Sun, & Ma, 2014). They are also important because different sensors have different characteristics, can make measures, by collecting measurement data in several types of conditions (e.g: weather conditions), and can be positioned on diverse positions, which ensures more accurate data (provides robust information about the environment).

Concluding, the transformation of a system based on centralized fusion to a system based on distributed fusion can be complicated, but has a lot of benefits to the sensor fusion process, since a sensor is a source and a deposit of information, the loss of only one component means that only one of the multiple channels of information has been lost (Khosla et al., 2017).
2.2.3 Types of Sensors

Before starting to list and present examples of sensors and the different types of data collected by them, it is relevant to define what a sensor is.

In their article, (X. Liu & Baiocchi, 2016) make a quite interesting survey about the characteristics of the sensors and smart sensors. In this way, (NIST, 2002) indicates that a sensor consists in a transducer whose main goal is to transform a physical, biological or chemical pattern into an electrical signal, collecting, in this way, digital data. The (Mark & Hufnagel, 2006) says that a smart sensor covers all the characteristics presented in the NIST description, plus a memory component. This memory component is responsible for storing sensor identification, calibration, correction of data, producer related information, and an interface to communicate. (Ahmad, 2016) specifies that a smart sensor, related to the IoT, has all the characteristics mentioned before, plus the integration of local computing capabilities allowing, in this sense, signal conditioning, data processing, data interpretation and decision-making process to happen in a local environment.

Through the last definitions, and in the context of this dissertation, we can point some key points that are relevant to understand what is a sensor, which are (Ahmad, 2016; Baskaran & Gangadharan, 2016; Mark & Hufnagel, 2006; NIST, 2002):

- Physical device;
- Interacts with physical stimulus (e.g: heat, light, among others);
- Conceives a specific impulse;
- Allows data storage and processing in a local environment.

As described in the abstract, the technological evolution and its importance have been creating more and more impact in the daily life of society. Thus, we are increasingly living in a technological world, where new technologies are appearing, sensors are no exception. Nowadays, sensors cover a huge variety of areas and purposes, so in this dissertation, only those that are related to the automotive industry, Bosch Car Multimedia Portugal S.A, and that give a good explanation about how sensors work will be presented.

Therefore, and according to the specifications presented previously, 2 sensors were identified, each one of them with its own characteristics, measures and purposes.

1. Camera Sensor (Web Cam)
2. Audio Sensor (Microphone)
2.2.4 Algorithms

Like all the others technological areas, the sensor fusion presents several algorithms that can be used to gather data from a multitude of sensors. Since there are a lot of methods and algorithms that are used in sensor fusion, only the most used, globally and in the automotive industry, will be discussed here.

The most part of the sensor fusion techniques are based on probability theories, mainly by the Bayes’ theorem. In this sense, it uses methods as Kalman filter (Standard, Extended, etc), Dempster-Shafer, sequential Monte Carlos methods, functional density estimates, among others (Baskaran & Gangadharan, 2016). Thus, Bayes’ theorem is the core of most of the data fusion methods (Durrant-Whyte & Henderson, 2016).

However, there are some Non-Bayesian based sensor fusion techniques that are also used. These techniques are based on interval methods, fuzzy logic, and the theory of evidence (Baskaran & Gangadharan, 2016). These non-probabilistic methods are proposed and used to deal with the limitations of the Bayesian based techniques, since the uncertainty is a great problem of the information fusion (Durrant-Whyte & Henderson, 2016).

In this sense, the limitations, or bottlenecks, of the probabilistic methods are (Durrant-Whyte & Henderson, 2016):

- **Complexity**: complication in defining a huge quantity of probabilities to apply the best probability based techniques;
- **Inconsistency**: difficulty in specifying a consistent set of probabilities and obtain a consistent output of the states of interest;
- **Models precision**: Necessity of being precise when specifying the probabilities for the events;
- **Uncertainty**: appears when there has lack of information derived from the sensors, or there is ignorance about them, leading, in this sense, to difficulties in specifying probabilities.

An exhaustive analysis of these algorithms can be observed next:

**(A) Bayes’ Theorem**

Bayes’ theorem involves conditional probabilities, these being quite important for this reason. This theorem consists of the probability of a given event happening, based on prior knowledge (Morettin & Oliveira Bussab, 2010).

Thus, the same is given by the following formula:

\[
(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(B)}
\]

where
According to the Bayes’ theorem for Multisensor Data Fusion, this requires conditional independence. In this sense, the multiplication of the likelihood function of each information source (new data collected from each sensor) and the prior probability gives rise to the posterior function taking into account all the observations $B_i$. This means that the likelihood probability of the information obtained from the different sensors and the prior information about $A$ can be encoded. However, the prior information of $B$ is not encoded, serving to normalize the posterior, for the validation of the model and for the association of the data (Durrant-Whyte & Henderson, 2016).

The previous observation can be analyzed by the following formula:

$$P(A|B^n) = P(A) \prod_{i=1}^{n} P(B_i|A)$$

(2)

where

- $n$: Number of sources of information (sensors)
- $P(A)$: Prior function ($A$ represents prior knowledge)
- $P(B_i|A)$: Likelihood function (new data)
- $P(A|B^n)$: Posterior function

The recursive expression of the Bayes’ theorem can be found next;

$$P(A_m|B^m) = \frac{P(B_m|A_m)P(A_m|B^{m-1})}{P(B_m|B^{m-1})}$$

(3)

where

- $A_m$: State vector that will be predicted at time $m$
- $B_m$: $A_m$ that will be observed at time $m$
- $A^m$: History of states
- $B^m$: History of observations

Through this formula, it’s only necessary to encode the posterior density ($P(A_m|B^{m-1})$), the one that contains all the prior information. In this sense, when the new data/information ($P(B_m|A_m)$) is collected, the posterior probability becomes the prior information and, after the its multiplication, followed by the normalization, it generates a new posterior density (Baskaran & Gangadharan, 2016).

Finally, the next state depends exclusively of the current state at time $m$ and can be given by:

$$P(A_m|B^{m-1}) = P(A_m|A_{m-1}) * P(A_{m-1}|B^{m-1})$$

(4)
Having completed this theorem, we can now move on to the techniques derived from it. In this sense, we will analyze the Kalman Filter, the sequential Monte Carlos methods, among others.

(a) Kalman Filter

The Kalman filter has more than 50 years but it is still one of the most used and, for this reason, most powerful sensor fusion algorithms for two reasons. The first one being the smooth noisy input data, and the second being the smooth estimate state (Souza, 2017). Thus, this algorithm is good with continued parametric states in automotive industry problems, like approximating localization, speed, among others. However, it is not the best option to estimate spacial occupancy, distinct labels, between others (Baskaran & Gangadharan, 2016).

In this sense, it represents a linear recursive statistical solution, consisting of a set of equations that allow to estimate, in an effective way, the state of a process that evolves over time (supports estimates of the process of the past, present and future of the states), with the minimum error variance when noisy data exists. Thus, it is a simple and intuitive concept with a high computational efficiency that allows solving the problem of discrete-data linear filtering, and helps to evaluate and track diverse features from each sensor (Baskaran & Gangadharan, 2016; Mau, 2005; Otto, 2013; Welch & Bishop, 2006).

It is a subset of the Bayesian filtering in which the density probability of states of a particular process follows a Normal distribution of $\mathcal{N}(\mu, \sigma)$. Through the following equations (these ones independent of each others), the Kalman filter as the main goal of estimate the state of a distinct-time controlled process (Welch & Bishop, 2006).

The first one consists on the state equation (process model) and is given by:

$$x_k = Ax_{k-1} + Bu_{k-1} + W_{k-1}$$

where

$x_k$ : State vector at time-step $k$
$A$ : State transition matrix
$B$ : Control-input matrix
$u_{k-1}$ : Control vector
$W_{k-1}$ : Process noise $v_k \sim \mathcal{N}(0, Q)$ with covariance $Q$ (changes over time-step, but we assume as constant)

The second one consists on the observation equation (measurement model) and is given by :

$$z_k = Hx_k + v_k$$

where
\[ v_k : \text{Observation/measurement noise } v_k \sim \mathcal{N}(0, R) \text{ with covariance } R \text{ (changes over time-step, but we assume as constant)} \]

\[ z_k : \text{State at moment } k \text{ and measurement noise linear combination} \]

\[ H : \text{Observation matrix} \]

Having completed the development of our model, it is necessary to evaluate and define the crucial parameters. In this sense, and through the initial information, we can identify two groups where the Kalman filter equations can be found (Figure 5). The first one consists of the time update (prediction) equations and the second one consists of the measurement update (correlation) equations (Baskaran & Gangadharan, 2016; Souza, 2017).

\[ z_k = A x_{k-1} + B u_k \]

\[ p_k = A p_{k-1} A' + Q \]

\[ x_k = x_{k-1} + K_k (z_k - H x_k) \]

\[ p_k = (1 - K_k H) p_k \]

**Figure 5:** Groups of Kalman filter equations (adapted from (Baskaran & Gangadharan, 2016; Souza, 2017))

To make this algorithm easier to understand, I will present a quick example where we can use it. As this pre-thesis is being developed in the automotive industry, let’s suppose we have some data (velocity data) collected from 3 GPS sensors (sensor used in this example). Through the formula presented below, this one that follows a normal distribution with mean \( \mu \) and standard deviation \( \sigma \), where the output consists on information used to develop a process using the
Kalman filter. In this sense $P(L_1) P(L_2)$ and $P(L_3)$ and the variance $\sigma_1^2$, $\sigma_2^2$ and $\sigma_3^2$, consist in information collected from sensor 1, 2 and 3 respectively (Baskaran & Gangadharan, 2016).

$$P(L) = \frac{1}{2\pi \sigma} e^{-\frac{(L-L_{\text{mean}})^2}{2\sigma^2}}$$  \hspace{1cm} (7)

where

$P(L)$ : Modeled information from sensor L

$\sigma$ : Standard Deviation

$L$ : Data derived from sensor L

Finally, the Kalman filter joins information through the calculation of measurements ($Z_{ik}$) and the equivalent kalman gain ($K_{ik}$) from different sources (n sensors) which the main goal is to update the state vector $x_k$ and the error covariance $p_k$ (Baskaran & Gangadharan, 2016).

**(b) Sequential Monte Carlo Methods**

The Sequential Monte Carlo (SMC) methods, also called Particle filter, are a group of methods developed based on simulations that offer an efficient, easy to implement and flexible approach to encode the posterior probabilities for non-Gaussian and nonlinear high dimensional data (Del Moral, 1996; Doucet, de Freitas, & Gordon, 2001).

In this sense, and as described before, these methods consist in a set of samples of a fundamental state space. After the creation of the samples, these methods, through the Bayes Theorem, simulate probabilistic inference. Thus, this process consists in repeat the sampling of the values to obtain the numerical results. This field has a lot of advantages nowadays, because of the advances in applied statistics, and his economical and good computational power (Doucet et al., 2001; Durrant-Whyte & Henderson, 2016).

The first step for the utilization of these methods is the creation of “n” number of samples at the time $k-1$, these ones that represent possible support points $e_{k-1}$ and attach his corresponding probability $p(e_{k-1})$ and normalized weights $w_{k-1}$.

The next steps can be analyzed on Figure 6:
Through the figure above, we start by predicting the support point $e_k$ at time $k$ through the last support point $e_{k-1}$, followed by the update of the probability density $p(e_k | z^k)$ and weights $w_k^{i}$. Finally, the exceptions (outliers) are removed resampling it through the weights, as we can see on "Outliers elimination" (Baskaran & Gangadharan, 2016; Durrant-Whyte & Henderson, 2016).

(B) Non-Bayesian Methods

As described in the introduction to this chapter, derived Bayesian algorithms and non-Bayesian algorithms are used in the fusion sensor process. In this sense, in this sub-section i will explain the usage of Dempster-Shafer, Fuzzy logic and interval methods on this field.

(a) Interval Calculus

Regarding the uncertainty, this method has advantages over probability-based techniques, since the definition of intervals helps in the measurement of uncertainty when insufficient probabilistic information is available. This technique leads with uncertain data through the definition of a lower and an upper bound, that is, let’s suppose that we have a state $x$ and that for that state we will define a lower (a) and an upper (b) bound. The statement of this interval will be given by $x \in [a, b]$ (Baskaran & Gangadharan, 2016; Durrant-Whyte & Henderson, 2016).

The definition of this interval doesn’t necessarily imply that the data follows a uniform distribution and, in addition, it contains the interval error that can be manipulated by simple and basic rules.

Let’s suppose now that we also have the bounds $a, b, c, d \in \mathbb{R}$, the data fusion interval (addition, subtraction, multiplication, and division) can be given by (where equations 8 and 10 are associative and commutative):

$$[a, b] + [c, d] = [a + c, b + d]$$  \hspace{1cm} (8)

$$[a, b] - [c, d] = [a - c, b - d]$$  \hspace{1cm} (9)

$$[a, b] \cdot [c, d] = \min(ac, ad, bc, dc), \max(ac, ad, bc, dc)$$  \hspace{1cm} (10)
It also allows to calculate and measure the metric distance. This calculation is given by:

\[ d([a, b], [c, d]) = \max(|a - c|, |b - d|) \] (12)

At last, it is possible the usage of arithmetic matrices. However, this process is not advised when the inversion of matrices is necessary. This method is not very helpful for sensor fusion techniques, since the need to establish the truth worry dependencies between the different sensors is complicated, making it difficult to merge data from them (Baskaran & Gangadharan, 2016; Durrant-Whyte & Henderson, 2016).

(b) Fuzzy Logic

The Fuzzy Logic method is often presented as an optimal tool for inexact reasoning, mainly to rules-based systems (Durrant-Whyte & Henderson, 2016). This method is characterized as a many-valued logic, through which the truth values of the variables can be any value (integer or decimal) between 0 and 1. Thus, it can contain the concept of partial truth, which means that it can go through a completely false variable (0) or a completely true variable (1) (Novak, Perfiljeva, & Mockor, 1999).

Next, I will describe the main operations of the Fuzzy logic methods. Let's suppose that we have a global set of elements \( X = \{ x \} \) and the sub-set \( Y \), this one with a specific characteristic and is contained in the first one \( Y \subseteq X \). In this sense, we can define a characterization/membership function \( \mu_Y(x) \), which main purpose is to report if a specific element \( x \in X \) is a member of the set \( Y \), as we can see on the next formula (Durrant-Whyte & Henderson, 2016):

\[
Y \mapsto \mu_Y(x) = \begin{cases} 
1 & \text{if } x \in Y \\
0 & \text{if } x \notin Y 
\end{cases} 
\] (13)

To make it easier to understand what was previously presented and considering the scope of this dissertation, let’s imagine that \( X \) is a set that contains all the existing cars and \( Y \) is a set that contains all the electric cars. This process is known as a crisp set. However, a fuzzy set is one through which the characterization function \( \mu_Y \) will define a degree of membership, between 0 and 1, as we can see on the next formula:

\[
\mu_Y(x) \rightarrow X[0, 1] 
\] (14)

Through the Figure 7 we can analyze the differences between the crisp and fuzzy sets.
2.2. Sensor Fusion

(a) Fuzzy Sets
\[ Y \ni \mu_Y(x) = \begin{cases} 1 & \text{if } x \text{ is totally in } Y \\ \leq \mu_Y(x) \leq 1 & \text{if } x \text{ is partly in } Y \\ 0 & \text{if } x \text{ is not in } Y \end{cases} \]

(b) Crisp Sets
\[ Y \ni \mu_Y(x) = \begin{cases} 1 & \text{if } x \in Y \\ 0 & \text{if } x \notin Y \end{cases} \]

Figure 7: Differences between fuzzy and crisp sets

Finally, we can observe that:
\[ A \cap B \ni \mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)], \quad (15) \]
\[ A \cup B \ni \mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)]. \quad (16) \]

Concluding, this method consists in determining the truth value of the variables (this value between 0 and 1), where 0 means completely false and 1 completely true. It can be also used the IF-THEN rule to determine this value.

(c) Dempster-Shater evidential reasoning technique

This theory can be translated into an alternative to traditional theories based on probabilities for the mathematical representation of uncertainty and imprecision (Sentz, Ferson, et al., 2002). It also analyzes whether the facts are valid or not based on beliefs, allowing, in this sense, the allocation of belief masses into sets or intervals (Baskaran & Gangadharan, 2016; Sentz et al., 2002).

Let’s suppose that we have a mutually exclusive/disjoint set of an event \( y = \{ \text{open}, \text{closed} \} \). If we were using a theory based on probabilities, we could assign a probability to each event (e.g. \( P(\text{open}) = 0.4 \), and thus \( P(\text{closed}) = 0.6 \)), but since we are using the Dempster-Shater evidential reasoning technique, we can create a set with all the subsets. This set can be analyzed on the next formula:

\[ 2^Y = \{ \{ \text{open}, \text{closed} \}, \{ \text{open} \}, \{ \text{closed} \} \} \]

(17)

to which belief masses are:
2.2. Sensor Fusion

\[ m(\text{open, closed}) : 0.6 \]
\[ m(\text{open}) : 0.2 \]
\[ m(\text{closed}) : 0.2 \]
\[ m(\Phi) : \emptyset, \text{belief mass 0 for normalization purposes} \]
\[ \sum_{E \in \mathcal{Y}} m[E] : 1, \text{contains all the } x \text{ possibilities of } E \text{ which } \sum = 1 \]

Analyzing the last formula, we can observe that there is a 60% of either open or closed, 20% chance of open, and 20% chance of closed. The first mass belief represents the inability or ignorance to distinguish between both alternatives.

Using the belief mass \( m \) for an evidence \( E \) we can determine the probability interval \( (P[E]) \). Let’s consider the next formula:

\[ \text{Bel}[E] \leq P[E] \leq \text{Pl}[E] \] (18)

To determine the belief measure of \( E \) \((\text{Bel}[E])\), this one that represents the lower bound, suppose that we have a subset \( B \) that is contained on \( E \). The sum of all the mass beliefs where \( E \) contains \( B \) represents the \( \text{Bel}[E] \). Thus:

\[ \text{Bel}[E] = \sum_{B \subseteq E} m[B], \quad \forall E \subseteq Y \] (19)

To determine the plausibility of \( E \), this one that represents the upper bound, the intersection of \( B \) with \( E \) must not be null. Then, the sum of all the mass belief of \( B \), that intersects with \( E \), represents \( \text{Pl}[E] \). Thus:

\[ \text{Pl}[E] = \sum_{B \in E \neq \Phi} m[B], \quad \forall E \subseteq Y \] (20)

or:

\[ \text{Pl}[E] = 1 - \text{Bel}[\sim E], \quad \forall E \subseteq Y \] (21)

Finally, to determine the probability interval \( P[E] \) it’s first necessary to determine the belief measure of \( E \) \((\text{Bel}[E])\) and the plausibility of \( E \) \((\text{Pl}[E])\).

Next, I will present an example where this method is used in a data fusion system perspective. In this systems, the \( \text{Bel}[E] \) determines how reliable a system is whereas \( \text{Pl}[E] \) represents the risk of a system. To fuse the evidences gather from each sensor we can use the following formula.

\[ m_{12} = \frac{1}{1 - K} \sum_{E \cap E' \neq \emptyset} m[E] m[E'] \] (22)
2.3 Big Data

where

$k$: conflict between the two evidences (if $k$ is null the evidences cannot be combined)

2.3 BIG DATA

2.3.1 Definition

With the constant increase in electronic devices, sensors, among others, more and more amounts of data are being generated, these with multiple degrees of complexity, at different velocities, from one or more sources. This significant increase of sources leads to the lack of conformity between the traditional tools and the amount of data generated and, consequently, to the inefficiency of these tools for the analysis or processing of those data (Chen, Mao, & Liu, 2014; Costa & Santos, 2016; Krishnan, 2013a; Zikopoulos & Eaton, 2011). It is also relevant to entitle datasets with a large amount of data (Bifet, 2013).

In this sense, and to specify what was presented previously, as a society and with the technological evolution, we are increasingly enjoying the technologies that are made available and, consequently, we generate more and more data.

This is where Big Data gains recognition, this term that first appeared in 1998 and was created by John Mashey for the different reasons presented above (Diebold, 2012).

In order to analyze and deal with all this data, new algorithms and tools are necessary. Thus, Doug Laney was the first person to introduce the 3 V’s when we talk about Big Data management, these ones that are very present in the definition and characterization of Big Data (Fan & Bifet, 2013). However, different authors have been proposing and trying to introduce new V’s to the concept of Big Data.

On Figure 8, we can find some examples of proposed V’s from different authors, where the black and the red ones represent, respectively, those that are defended by the authors and those that are not defended by them.
Regarding the different V’s presented we can indicate that:

- **Volume**: Consists of increasing the amount of data that is generated, which have reached unprecedented proportions. It also represents an increase of the data sources, such as data from sensor networks, nuclear plants, aircraft engines, among others (Fan & Bifet, 2013);

- **Velocity**: Represents the velocity at which data is collected and made available in order to obtain useful information, that is, the velocity of data processing (Fan & Bifet, 2013);

- **Variety**: Characterizes the different types of data, such as sensory data, audio, video, meteorology, among others, collected from heterogeneous data sources (Fan & Bifet, 2013). These data can be structured, semi-structured and unstructured (Simon, 2013);

- **Variability**: Consists of changes in data structures and how users want to interpret that data (Fan & Bifet, 2013);

- **Veracity**: The data sources that provide the data present different degrees of quality, regarding the coverage, precision, and timeliness of the provided data (Luna Dong & Srivastava, 2013);

- **Value**: Represents the value of the data for a given business providing, in this sense, a competitive advantage in the decision-making process (allows answers to previously defined questions) (Fan & Bifet, 2013);
- **Volutility**: Characterizes how long the data is valid and how much time it should be stored. In this sense, and since we live in a world where data is collected in real-time, it is necessary to analyze the usefulness of certain data for the current business. If this data is irrelevant it must be deleted (u. d. Khan et al., 2014).

The main challenge related to this area is the ability to exploit large volumes of data and extract information that is relevant to future decisions in a given business context (Leskovec, Rajaraman, & Ullman, 2014). This process often has to be done very efficiently and closely of reality due to the inability to store all observed data.

Through the analysis of the three major V's for Big Data, we can find a high complexity regarding this process (Krishnan, 2013a). Thus, we also verify that there exists (Figure 9) (Krishnan, 2013a):

- **Ambiguity**: The lack of metadata leads to ambiguity;
- **Viscosity**: Measures the resistance, that is, the decrease of the velocity of the flow in the volume of data, being able to be found in datasets, business rules, among others;
- **Virality**: Measures and describes how quickly data is shared between people through a peer-to-peer network.

(Wu, Zhu, Wu, & Ding, 2014) also propose a theorem called HACE to represent the different characteristics of Big Data. Thus, the characteristics present in this theorem are (Wu et al., 2014):

- **Heterogeneous**: Consists of large amounts of data represented by diverse and heterogeneous dimensionalities, since different data sources use their own schemata to record data;
2.3. Big Data

- **Autonomous**: The data sources are autonomous with decentralized and distributed controls. In this sense, each source is capable of generating and collecting data without using centralized controls;

- **Complex**: Since the volume of Big Data is increasing, the complexity and the relationship between data are also increasing;

- **Evolving**: Big Data takes into account the complexity of the relationships between the data, along with any changes that may occur, in order to determine useful information.

Concluding, Big Data can be defined as a large amount of data (volume) made available with various levels of complexity (variety), generated at different velocities (velocity) and with different levels of ambiguity, that cannot be processed using traditional technologies, methods, and algorithms. In this sense, it requires techniques, algorithms, and methods (such as NoSQL databases and Hadoop) that are most effective for capturing, managing and analyzing high-speed data. These methods result in a significant cost reduction and greater scalability than the previously used methods (Krishnan, 2013a).

2.3.2 Types of NoSQL Databases

In this chapter, before starting to enumerate the best NoSQL databases that can be used on this dissertation, it is relevant to indicate what NoSQL is, the CAP theorem and its characteristics and the different data models that exists to classify those databases.

In this sense, and with all the arguments presented before to the appearance of Big Data, the relational databases (RDBMS) where not capable to support the necessary scalability requirements for huge amounts of transactional data, such as applications like sensors networks, web applications, among others (Krishnan, 2013b). For this, it was necessary to develop data could process Big Data efficiently, ensuring good reading and writing performance (J. Han, E, et al., 2011).

These new databases should be able to support high-speed data requirements for various types of applications (high-concurrency and large scale applications) in an environment that was not completely SQL (J. Han, E, et al., 2011; Krishnan, 2013b).

This is where NoSQL and the NoSQL databases are born. The NoSQL, also known as “Not Only SQL” and “Not Relational” (Cattell, 2011) has emerged, like its name implies, as a need for non-relational and non-SQL-driven databases, this ones that work with similar architectures but with different workloads (Krishnan, 2013b).

Thus, the NoSQL presents several categories of NoSQL databases, each one with its specific attributes, advantages and limitations, given that there is no single solution that is better than the others. However, there are certain NoSQL databases that are better to solve specific problems (Rai & Chettri, 2017).
To classify these databases, (Brewer, 2010) proposed, in 2000, the currently famous CAP theorem. This theorem presents 3 requirements that are relevant to classify systems (e.g. databases). These requirements are (Krishnan, 2013b):

- **Consistency**: All the data is available in all the nodes or systems;
- **Availability**: Every request will have a response;
- **Partition Tolerance**: The system will continue its function regardless of failures.

![CAP Theorem](image)

The main objective of this theorem indicates that a distributed system cannot make use of the three requirements simultaneously, but only two. Thus, we can assign three classifications to the NoSQL databases through this model. The three classifications are (J. Han, E, et al., 2011):

- **Databases concerned with the consistency and the availability (CA)**: They present partition tolerance problems and typically uses replication to ensure data availability and consistency;
- **Databases concerned with the consistency and the partition tolerance (CP)**: They have problems with data availability, however the data stored in distributed nodes and they guarantee the consistency of the same;
- **Databases concerned with the availability and the partition tolerance (AP)**: They achieve "eventual consistency" and guarantee data availability and partition tolerance.

Finally, we can also classify the NoSQL databases through their data model. In this sense, most of the NoSQL databases differ from relational ones by their data models (Funck & Jablonski, 2011). Thus, these systems can be classified into four categories:
• **Key-value stores:** This model is implemented based on a hash table that has a unique key and a pointer to a specific value of data, creating a key-value pair (Krishnan, 2013b). Concluding, this model means that a determined value corresponds to a determined key (J. Han, E, et al., 2011) and a database with this model is a schemaless collection of key-pairs which presents data access operations in individual or groups of key pairs (Atzeni, Bugiotti, Cabibbo, & Torlone, 2016);

• **Document stores:** This data model and the key-value pair are very similar at the structure level. Like its name implies, databases that use this data model are designed to manage and store documents (Moniruzzaman & Hossain, 2013). The data is stored as a document and represented in several formats, like JSON, XML, PDF, among others (Nayak, Poriya, & Poojary, 2013). It is also very flexible because allows listing multiple levels of key-pair values (Krishnan, 2013b);

• **Graph databases:** Data model through which databases support scalability across a cluster of machines. It is also based on graph theory (Krishnan, 2013b);

• **Wide Column (or Column Family) stores:** This model represents a distributed, column-oriented data structure, that can accommodate multiple attributes per key (Moniruzzaman & Hossain, 2013). Although using tables, does not allow the association between them (J. Han, E, et al., 2011).

After all this analysis, we are able to describe more specifically the different NoSQL databases. Currently, there are more than 225 NoSQL databases (NoSQL, n.d.). However, only those that are more referenced in literature and that can be used in this dissertation will be addressed.
<table>
<thead>
<tr>
<th>Database</th>
<th>CAP</th>
<th>Data Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBASE</td>
<td>CP</td>
<td>Wide Column stores</td>
<td>HBase represents a NoSQL database that was developed over the HDFS, with high scalability and distributed fault tolerance. In addition, it is a non-relational, multi-dimensional and open-source database that was developed in JAVA (Cattell, 2011). It also guarantees high availability and performance in order to support the storage and processing of data sets (Krishnan, 2013b) and be used for real-time read and write random access. As for its model, and as explained above, it is column-oriented, in which the data is organized into labeled tables. Each table is stored with a multidimensional sparse map, this with columns and rows. Thus, each row has a sorting key and an arbitrary number of columns. In conclusion, HBase is a scalable, distributed database that enables structured data storage for large tables (Krishnan, 2013b)</td>
</tr>
<tr>
<td>HIVE</td>
<td>-</td>
<td>-</td>
<td>HIVE represents an open-source data warehousing solution developed at the top of Hadoop to enable summarizing, querying, and analyzing data (Venner, 2009). HIVE supports the Hive QL language to organize, aggregate and run queries on the data. Hive QL is very similar to SQL, using a declarative programming model (Ghaffar Shoro &amp; Soomro, 2015).</td>
</tr>
<tr>
<td>Cassandra</td>
<td>AP</td>
<td>Wide Column stores</td>
<td>Cassandra is an open-source NoSQL database that was created by Facebook and developed in JAVA (J. Han, E, et al., 2011; Nayak et al., 2013). It consists of a scalable multi-master database without any single point of failure (Krishnan, 2013b), this is, if a part of the system fails, the entire system won’t stop working (Dooley, 2001). This database has good characteristics as the fact that his scheme is very flexible and does not need to do the scheme first, the capacity to support range queries (it can range queries for key), to have high scalability, that is, the loss of a certain component does not affect an entire cluster, and supports linear expansion (J. Han, E, et al., 2011). Concluding, Cassandra consists of a distributed storage system that allows managing huge amounts of structured data through commodity servers and guarantees high availability, partition tolerance, persistence, among others (Lakshman &amp; Malik, 2010). As for disadvantages, this database reads are slower than writes (Nayak et al., 2013).</td>
</tr>
</tbody>
</table>

(Table 2 continued on next page)
(Table 2 continued from previous page)

<table>
<thead>
<tr>
<th>Database</th>
<th>CAP</th>
<th>Data Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MongoDB</td>
<td>CP</td>
<td>Document Store</td>
<td>MongoDB consists of a database located between a relational database and a non-relational database (J. Han, E, et al., 2011) that was developed through C++. It is efficient and with high performance (Nayak et al., 2013). Its functionalities allow consistency, fault tolerance, persistence data aggregation, fast access to huge amounts of data, ad hoc queries, indexing, among others (J. Han, E, et al., 2011; Nayak et al., 2013). Finally, supports BSON (Binary JSON) data structures to store complex data types (Cattell, 2011). BSON documents contain an ordered list of elements, such as the name, type, and value of a field (Nayak et al., 2013).</td>
</tr>
</tbody>
</table>

### 2.4 MACHINE LEARNING

#### 2.4.1 Definition

Machine Learning has reached tremendous amounts of progress in the last few decades, with a high number of systems using its algorithms in fields like medicine, industry, among others (Kononenko & Kukar, 2007).

The main drive of this area is to give systems/machines the ability to learn diverse characteristics without the need to be programmed (Simon, 2013), allowing them to learn from previous experiences to optimize their performance (Sheng & Zhou, 2017).

As so, Machine Learning is a field of computer sciences that gives computers the ability to learn without being explicitly programmed and tries to create algorithms that can be applied to a set of existing data, helping machines to find standards on the information that correlate to a specific result (Stanton, 2013), meaning that its main worry is to find models, standards and other regularities in the data (Fürnkranz, Gamberger, & Lavrač, 2012). It is a branch of Artificial Intelligence (Mohri, Rostamizadeh, & Talwalkar, 2012).

Most of the times, Machine Learning is used to extract useful information from data and organize it in an understandable way, granting a better understanding of this same data, allowing it to be used for a distinct number of purposes. (Witten, Frank, Hall, & Pal, 2016).

It has become common to find technologies that make use of machine learning algorithms, such as information filtering that is commonly presented to social network users, among others. Furthermore, there has been a growth in the number of applications that use these algorithms and they paved the way to develop processes like object identification, voice-to-text transformation, among others.
When we talk about Machine Learning models, they can assume various types. These types can be identified as follow:

- **Supervised Learning**: Model with a set of labelled data with the correct answers through which the machine can learn (Goodfellow, Bengio, & Courville, 2016; Grus, 2015). It assumes that the training examples are classified (J. Han, Pei, & Kamber, 2011).

- **Unsupervised Learning**: Model where the data is not labelled meaning that the machine does not learn from correct data. It worries itself with the analysis of non-classified examples (Goodfellow et al., 2016; Grus, 2015);

- **Semi-supervised Learning**: Model where some of the data has labels and the rest of it is unlabelled (Goodfellow et al., 2016; Grus, 2015);

- **Reinforcement Learning**: Model applied to dynamic environments and to which is provided a reward signal that tells the algorithm if he is performing well. Afterwards the algorithm must learn the best course of action (Goodfellow et al., 2016);

- **Online Models**: in which the model needs to continuously adjust to newly arriving data;

- **Active Learning**: Allows the users to have a preponderant and active role in the learning process, with the goal of optimizing the quality of a model through the active acquisition of knowledge from the users (J. Han, Pei, & Kamber, 2011).

These were all the identified Machine Learning models, with most authors complementing each other, in other words, (Grus, 2015) presents the concepts of Supervised Learning, Unsupervised Learning, Semi-Supervised Learning and Online Models. (Goodfellow et al., 2016) presents the concept of Reinforcement Learning and, finally, (J. Han, Pei, & Kamber, 2011) presents a new way of learning, known as Active Learning.

Finally, Machine Learning is usually divided into Supervised Learning e Unsupervised Learning (Goodfellow et al., 2016).

Like all areas, Machine Learning presents some problems that may compromise the efficient development of a project. This is where the concepts of Overfitting and Underfitting come in. So:

- **Overfitting**: consists in creating a model that performs well on data that has been trained upon. However, it generalizes poorly on new data (Grus, 2015). Imagine that we have a dataset with people with different facial expressions and that we train a model with this dataset. Then, we tested the model with the original dataset and it predicts results with 99% accuracy. Finally, we run the model in a new dataset with facial expressions and we only got 50% accuracy. This example represents a case of the occurrence of overfitting;

- **Underfitting**: covers the creation of a model that does not perform well even on training data. Usually, when this happens, it is decided that the model is not good enough and the search for
a better one continues (Grus, 2015). Imagine again that we have a dataset with various facial expressions, and that we train a model with this dataset. We then tested the model with the original dataset and it predicts results with 40% accuracy. This example represents a case of the occurrence of underfitting.

Finally, we can talk about the Bias-Variance Tradeoff. Bias and Variance are measures of what would happen if we retrained our model many times on different sets of training data. If our model has high bias (performs poorly even on the training data) and low variance (since the use of two different training sets leads to very similar average values), it corresponds to underfitting. If our model has high variance (since the use of two different training sets leads to very different models) and low bias (fit the training set perfectly), it corresponds to overfitting (Grus, 2015).

![Bias-Variance dilemma](Figure 11: Bias-Variance dilemma (adapted from (Fortmann-Roe, 2012))

### 2.4.2 Classification

Machine Learning presents several algorithms for the most diverse purposes, for example, regression, classification, segmentation, among others. In this dissertation, the classification algorithms will be the main focus, since the dissertation's main purpose is to correctly identify people’s emotional patterns.

For a better understanding of this dissertation approach the scheme on Figure 12 was created. This figure illustrates the process used for the Machine Learning component.
According to (Domingos, 2012), there are several types of Machine Learning. However, the most mature and widely used is the classification one. A classifier represents a system that inputs a vector of continuous and/or discrete feature values and outputs a single discrete value: the class. Thus, classification extracts models that describe important data classes, and predicts categorical (discrete, unordered) class labels (J. Han, Pei, & Kamber, 2011).

A lot of classification methods have been developed and proposed by researchers in Machine Learning, pattern recognition and statistics. These classification methods have different purposes, like fraud detection, target marketing, among others (J. Han, Pei, & Kamber, 2011).

In a more technical way, we can say that a classification model has the ability to automatically learn the implicit classification function via a training process. After the training process, the resulting model can be used to predict, with high accuracy, the class of “unseen” values (new elements that need to be classified) (Domingos, 2012).

A classification model depends on (Domingos, 2012):

- the training set;
- knowledge representation;
- the machine learning algorithm.

The three main challenges of the classification models are:

- **Accuracy**: represents the percentage of test set data that are correctly classified by the classifier. The associated test for each label of each test is compared with the learned class prediction for that test and, if the accuracy of the classifier is considered acceptable, it can be used for future classifications (C. Z. Han et al., 2006). Thus, the accuracy is considered one of the most important classification algorithms (Maimon & Rokach, 2005).

- **Efficiency**: the creation of a classifier can be difficult from the point of view of processing (Maimon & Rokach, 2005);
- **Easy to Use**: a classification model must be easy to use in applications (Maimon & Rokach, 2005).

In 2008, (Wu et al., 2008) reported the top 10 best algorithms. Among these algorithms, there are 6 classifiers, including SVM, KNN, AB, NB, CART and C4.5.

Also, due to the increasing scale and complex nature of many real-world problems in different areas, new classifiers are continuously being proposed (Zhang, Liu, Zhang, & Almpanidis, 2017).

Finally, the classification process consists in two-steps: the learning step (where the classification model is developed) and the classification step (where the model must predict the different class labels for a given data) (C. Z. Han et al., 2006).

In the first step, a classifier is built with a predetermined set of classes or concepts. This consists on the learning step (or training phase), where a classification algorithm builds the classifier by analyzing or learning from a training set and their associated class labels (C. Z. Han et al., 2006).

The first step of the classification process can be the learning of a function that is able to separate the data classes. The second step consists in using the model for classification, where the predictive accuracy of the classifier is estimated through the use of a test set (independent from the training set, which means that they were not used to build the classifier) (C. Z. Han et al., 2006).

### 2.4.3 Neural Networks

An ANN, the abbreviation of Artificial Neural Network, is a predictive model, whose performance is based on the way the brain operates. The brain consists of a collection of neurons wired together, where each of the neurons looks at the outputs of the other neurons that feed into it. After that, it does a calculation and then either fires (if some threshold is exceeded) or it doesn’t (if the thresholds are not exceeded) (Grus, 2015).

The ANNs present artificial neurons, which perform similar calculations to the human brain over their inputs. They are designed to cover and solve a wide variety of problems, such as facial detection and classification, and are often used on Deep Learning (Grus, 2015).

The Neural Networks operation is similar to the human brain because the knowledge is obtained from the environment and stored on its connections, through a learning process, that is, they learn with the experience of past data patterns (Haykin et al., 2009).

Thus, a Neural Network represents many simple connected processors, commonly called neurons. Each neuron produces a sequence of real-valued activations, where the input neuron is activated through sensors that perceive the environment, and the other neurons are activated through weighted connections from the previously active neuron (Schmidhuber, 2015).

Some of this neurons may influence the environment by triggering some actions. The learning assignment consists in finding weights that make the NN demonstrate the desired behavior (Schmidhuber, 2015), such as emotion classification or driving a car. Finally, depending on the problem and the connections between neurons, such behavior may require long and causal chains of computational
2.4. Machine Learning

stages, where each stage transforms the aggregate activation of the network, normally in a non-linear way (Schmidhuber, 2015).

On Figure 13, we can observe a simplified example of an Artificial Neural Network.

![Artificial Neural Network with one hidden layer](image)

Figure 13: Artificial Neural Network with one hidden layer

When we talk about classification, the ANNs have presented good results on accuracy and prevision if compared to other models (Cortez, 2010). However, most of the Neural Networks are like “black boxes” because if we inspect their details, we won’t understand how they are trying to solve certain problems. They are also very difficult to train when we have a large Neural Network (Grus, 2015).

Thus, we can say that the ANNs are mathematical models that were developed to mimic the storing and processing capabilities of some information of the brain, and are widely used on recognition and prediction patterns (Yu, Hu, & Bao, 2003).

There are several types of Neural Networks. These Neural Networks and its characteristics can be found on Figure 14.
Finally, we can talk about two learning procedures that are used on Neural Networks, such as:

- **Back-propagation**: consists on a method used in Artificial Neural Networks to calculate the gradient needed in the calculations of the weights that are going to be used in a network (Goodfellow et al., 2016). It is commonly used on deep neural networks (Nielsen, 2015). This procedure continuously adjusts the weights of the connections with the main purpose of minimizing the difference or error between the actual and desired output vector of the net. Finally, as a result of the weight adjustments, the internal ‘hidden’ units, these units that are not part of the input or output units start to represent important features of the task domain, these being regularities from that task captured by the interactions of these units (Rumelhart, Hinton, & Williams, 1986);

- **Feed-forward propagation**: consists in an artificial neural network where the connections between the different units don’t form a cycle (Zell, 1994). In this sense, it represents discrete layers of neurons where each of them is connected to the next. Normally, it entails an input layer that receives inputs and feeds them forward, one or more ‘hidden layers’ (each one consists of neurons that use the outputs of the previous layer to perform some calculations and pass the result to the next layer), and an output layer that produces the final outputs. Each non-input neuron has a weight corresponding to each of its inputs and bias (Grus, 2015).

---

**2.5 SOTA Relevant Sections**

For the development of this state of the art, it was necessary to carry out an in-depth research of the various concepts related to this dissertation, as we can see in table 3.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Area of Interest</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boström et al. 2017</td>
<td>Sensor Fusion</td>
<td>(Boström et al., 2017) make a very interesting survey of the different definitions of fusion sensor in a chronological way.</td>
</tr>
<tr>
<td></td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>Sun et al. 2017</td>
<td>Sensor Fusion</td>
<td>(Sun et al., 2017) make a very brief description of two architectures used to carry out the Sensor Fusion process, and its advantages and disadvantages.</td>
</tr>
<tr>
<td></td>
<td>Architectures</td>
<td></td>
</tr>
<tr>
<td>X. Liu and Baiocchi 2016</td>
<td>Sensors and Smart Sensors</td>
<td>(X. Liu &amp; Baiocchi, 2016) present a bibliographical research on the characteristics of sensors and smart sensors.</td>
</tr>
<tr>
<td>NIST 2002</td>
<td></td>
<td>(NIST, 2002) presents a definition of what a sensor is and its characteristics.</td>
</tr>
<tr>
<td>Mark and Hufnagel 2006</td>
<td></td>
<td>(Mark &amp; Hufnagel, 2006) presents a definition of what a smart sensor is and its characteristics.</td>
</tr>
<tr>
<td>Ahmad 2016</td>
<td></td>
<td>(Ahmad, 2016) presents a definition of what a smart sensor is and its characteristics when related to the IoT.</td>
</tr>
<tr>
<td>Baskaran and Gangadharan 2016</td>
<td>Sensor Fusion Algorithms</td>
<td>(Baskaran &amp; Gangadharan, 2016) present a rather in-depth description of the most widely used fusion sensor algorithms in the automotive industry.</td>
</tr>
<tr>
<td>Krishnan 2013a</td>
<td>Big Data</td>
<td>(Krishnan, 2013a) makes a very explicit description of what big data is, what it consists of, and the need for it.</td>
</tr>
<tr>
<td>Wu et al. 2014</td>
<td>Big Data Characterization through HACE theorem</td>
<td>(Wu et al., 2014) proposes a theorem to classify Big Data. This theorem is called HACE.</td>
</tr>
</tbody>
</table>

(Table 3 continued on next page)
### Table 3

<table>
<thead>
<tr>
<th>Reference</th>
<th>Area of Interest</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan and Bifet 2013</td>
<td></td>
<td>(Fan &amp; Bifet, 2013) indicate that, in 2000, Doug Laney proposed a characterization of Big Data through 3 V's and propose a characterization through 5 V's.</td>
</tr>
<tr>
<td>Sagiroglu and Sinanc 2013</td>
<td></td>
<td>(Sagiroglu &amp; Sinanc, 2013) propose a characterization of Big Data based on 3 V's.</td>
</tr>
<tr>
<td>Luna Dong and Srivastava 2013</td>
<td>Big Data Characterization through V's</td>
<td>(Luna Dong &amp; Srivastava, 2013) propose a characterization of Big Data based on 4 V's.</td>
</tr>
<tr>
<td>Sharma 2015</td>
<td></td>
<td>(d. Almeida &amp; Bernardino, 2015) propose a characterization of Big Data based on 5 V's.</td>
</tr>
<tr>
<td>Krishnan 2013b</td>
<td>NoSQL Definition</td>
<td>(Krishnan, 2013b; Rai &amp; Chettri, 2017) present the concept of NoSQL and why it needs to be used.</td>
</tr>
<tr>
<td>Rai and Chettri 2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. Han, E, et al. 2011</td>
<td>NoSQL databases characterization</td>
<td>(J. Han, E, et al., 2011) presents a brief description of the CAP theorem and the NoSQL databases characterization through it.</td>
</tr>
<tr>
<td>Krishnan 2013b</td>
<td></td>
<td>(Krishnan, 2013b) presents a brief description of the NoSQL databases characterization through the data model.</td>
</tr>
<tr>
<td>Grus 2015</td>
<td>Machine Learning Definition</td>
<td>(Grus, 2015) and (Goodfellow et al., 2016) present a good description about what is Machine Learning, the different types of Machine Learning and problems related to this area.</td>
</tr>
<tr>
<td>Goodfellow et al. 2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domingos 2012</td>
<td>Machine Learning Classification</td>
<td>(Domingos, 2012) and (Maimon &amp; Rokach, 2005) present a good description about the type of Machine Learning named Classification.</td>
</tr>
<tr>
<td>Maimon and Rokach 2005</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Table 3 continued on next page)
### Table 3 continued from previous page

<table>
<thead>
<tr>
<th>Reference</th>
<th>Area of Interest</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grus 2015 Haykin et al. 2009</td>
<td>Neural Networks</td>
<td>(Grus, 2015) and (Haykin et al., 2009) talk a little bit about what is a neural network and their characteristics.</td>
</tr>
</tbody>
</table>

| Z.-T. Liu et al. 2018 | | (Z.-T. Liu et al., 2018) presents a method to classify emotions from speech based on decision trees. |
| Tzirakis, Trigeorgis, Nicolaou, Schuller, and Zafeiriou 2017 | | (Tzirakis et al., 2017) presents an multimodal emotion recognition system based on Deep Neural Networks. |
| Bahreini, Nadolski, and Westera 2016 | | (Bahreini et al., 2016) presents a framework for real time emotion recognition. |
In Chapter 2, the concepts relevant to this dissertation were reviewed, such as Sensor Fusion, Big Data and Machine Learning. In this sense, the practical cases on which this dissertation was initially based and that will be addressed are presented below.

- **Case Study 1**: This project aims to analyze a database of geo location data of city buses fleet data and characterize city traffic flow. At a second step augment the database with contextual information and improve the characterization of city traffic flow by applying data mining techniques. Finally, develop a model for acquisition of new data considering both the interior and exterior sensing opportunities in vehicles/buses.

- **Case Study 2**: This project seeks the development of a system that can collect synchronized audio and video data within a vehicle, followed by the use of this data to develop a platform capable of recognizing a person’s emotional state through facial and vocal gestures. To do so, the ROS (Robot Operating System) will be used as a data synchronization method and a machine learning project will be developed for each type of data, taking advantage of its various algorithms. Finally, this project will be finalized with the fusion of the models created for each one.

With these two case studies in mind, it was decided, together with the advisor, that the case study to be explored under this dissertation was the second one, since the first case presents inconsistencies in the data collected.

For this practical case, as it was already described, the ROS will be used to synchronize several sensors and collect data, to facilitate the process of labeling them.

Then, two machine learning models will be developed, one for the audio data and the other for the visual data. For that, different algorithms will be analyzed, to analyze which ones obtain the best accuracy during the emotional classification of a person. In order to make the system more robust, different people will be used to create the data sets, thus avoiding the overfit of the models.

Several cars will also be used since the project is being developed from a car industry perspective, that is, only the environmental characteristics of the cars are relevant. Data augmentation concept will also be used, such as rotating an image.

Finally, these models must be fused to make a better representation of the system.
WORKPLAN

To ensure that the initially defined objectives will be successfully achieved without compromising the project development, its tasks must be carefully planned. For that, a set of deadlines were established, that must be fulfilled and used as a basis to manage, control and plan the different activities and stages of the project evolution. This planning will give a broad view of the project and will ensure that no task takes longer than the expected time to carry out.

Moreover, this dissertation presents some complexity and, for this reason, it has some risks associated with it. These risks need to be managed and controlled so they cannot compromise the development and success of this project, that is, if in the beginning we make an efficient planning, we can anticipate problems and solve them quickly. Two risks that must be taken into account are: the possible change of the tasks / requirements and the failure to meet the deadline attributed to each task.

In table 4, we can find all the deadlines defined since the beginning of the project until the end of it. These deadlines were used as a basis for the project planning development.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work plan submission</td>
<td>until October 2, 2017</td>
</tr>
<tr>
<td>Dissertation project delivery</td>
<td>until January 22, 2018 or February 19, 2018</td>
</tr>
<tr>
<td>Dissertation project evaluation</td>
<td>until February 5, 2018 or March 5, 2018</td>
</tr>
<tr>
<td>Dissertation document delivery</td>
<td>until June 22, 2018 or October 22, 2018</td>
</tr>
<tr>
<td>Public proofs</td>
<td>until December 21, 2018</td>
</tr>
<tr>
<td>Delivery of the final version of the dissertation document</td>
<td>until December 28, 2018</td>
</tr>
</tbody>
</table>

The previous table allows the characterization of the main tasks for the dissertation's work plan. The nature of these tasks can be of two different types: theoretical component tasks and practical component tasks. The theoretical component tasks consist in the search, collection, and understanding
of all the concepts considered relevant and fundamental for the practical development. In turn, the practical component tasks concern all the steps required for an efficient functional development of the system in question.

Table 5: Dissertation’s Work Plan

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Duration</th>
<th>Brief description of the tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>One month</td>
<td>Development and subsequent revision and delivery of the work plan, which contains information about the topic, abstract, objectives and expected results, schedule and, finally, the methodological approach of this dissertation.</td>
</tr>
<tr>
<td>Task 2</td>
<td>Three months and a week</td>
<td>Initial analysis of the TUB’s database, followed by the development of the data dictionary and the search for inconsistencies and failures in the data. Loading the database into a NoSQL database using the ETL process, finishing with a verification of the data to know if it is useful for the development of the project.</td>
</tr>
<tr>
<td>Task 3</td>
<td>Three months</td>
<td>Development of a literature review about the concepts relevant to this dissertation, followed by the analysis of the work done so far related with the subject. Elaboration of the abstract and scope, identification of the objectives, motivations, expected results and problem statement of the project, followed by the elaboration of the methodological approach and scheduling of the tasks. Finally, revision and correction of the elaborated documentation and its delivery (pre-dissertation).</td>
</tr>
<tr>
<td>Task 4</td>
<td>One week</td>
<td>Discussion of the relevance of the database for the dissertation, taking into account the analyzed data. Introduction of a new proof of concept that consists in the recognition of the emotions of the drivers, through sensor fusion and machine learning techniques.</td>
</tr>
<tr>
<td>Task 5</td>
<td>Two weeks</td>
<td>Acquisition of knowledge regarding the operation of the tools being used, through tutorials available online. These tutorials have information about how to synchronize different nodes (sensors), record and call back data, use python and C++ to develop subscribers and publishers, among others.</td>
</tr>
</tbody>
</table>

(Table 5 continued on next page)
Develop a sensor fusion system that, in the beginning, synchronizes and collects data of two sensors (camera and audio). For this task it will be used the Robot Operating System (ROS), since this framework represents an easy way calibrate and synchronize different types of sensors. This framework will help in the labeling process since we can analyze the emotional state in a more specific way, that is, we can get a better perception of a person emotional state through her facial and speech expressions.

Acquisition of two datasets for training and testing. This process involves the synchronization of the two sensors (camera and audio) in which different people from both genders with different facial expressions and voice tones should be used. The goal of this heterogeneity is to ensure a better efficiency in the emotional recognition of the human beings.

Development of a machine learning project that allows the use of its algorithms for classification and detection purposes. This machine learning project should lead to the development of two models, one to classify emotions through images, and the other to classify emotions through sound.

Analysis and testing of several machine learning algorithms with the objective of determine the best ones for the different types of data. In this sense, we can analyze the models that obtained better final results, such as accuracy.

Use of data augmentation to increase the volume and diversity of the datasets (audio and image) to be used for the purpose of improving the results obtained.

Fusion of the two algorithms developed to obtain a better classification of a person’s emotional state. This process will reduce the level of uncertainty related to the classified emotion.

Writing and review of the dissertation report.
(Table 5 continued from previous page)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Duration</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 13</td>
<td>One day (October 31, 2018)</td>
<td>Delivery of the dissertation report</td>
</tr>
<tr>
<td>Task 14</td>
<td>One week</td>
<td>Preparation of the presentation of the dissertation.</td>
</tr>
</tbody>
</table>

Finally, in Figure 15 we can find all the previously described tasks and their respective deadlines, up to the final deadline of the project.

![Figure 15: Gantt Diagram with all the tasks](image-url)
REFERENCES


Atzeni, P., Bugiotti, F., Cabibbo, L., & Torlone, R. (2016, 10). Data modeling in the nosql world.


References


Mau, S. (2005, 12). What is the kalman filter and how can it be user for data fusion?


References


This thesis was written under a curricular internship at Bosch Car Multimedia Portugal S.A.