UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTÀ DI INGEGNERIA

Corso di Laurea Magistrale in Ingegneria Meccanica

Wind noise quality assessment in full vehicle End-of-Line NVH testing adopting neural networks

Utilizzo di reti neurali per la valutazione della qualità del rumore dovuto all'aerodinamica di un autoveicolo nei test NVH a fine ciclo di produzione

Relatore:

Prof. Paolo Castellini

Correlatore:

Dott. Claudio Colangeli

Tesi di:

Mara Magnoni

Anno Accademico 2019/2020

Table of Contents

List of Figures

Abstract

In recent times, with the growing popularity of electric and hybrid vehicles, the interior acoustic comfort has become one of the main key factors in vehicle development process with the aim of reducing the level of interior noise and improving the sound quality as well. As engine and road-tire noise has been minimized and as vehicles speed has increased over the last years, aerodynamic noise has become an important factor that influences the interior acoustic comfort. The two main wind noises that determine discomfort and annoyance in a vehicle are leakages and whistling noises. At the time of writing, they are detected and analyzed in a full vehicle End-of-Line using purely subjective methods. However, with the recent advances and transformative potential of Machine Learning (ML), the Neural Networks can be used to reduce vehicle evaluation time and sound engineers' efforts when compared with standard acoustic NVH (Noise, Vibration and Harshness) optimization procedures.

The purpose of this thesis, made in collaboration with *Siemens PLM Software,* is to develop Neural Networks (NNs) for the recognition of leakages and whistling noises in a full vehicle End-of-Line. The study starts with the database generation through an executable digital twin of a virtual end-of-line vehicle, where the Vehicle Sound Simulator methodology and the procedure for the creation of the NN databases are described. The second part of the thesis focuses on the NN development with an emphasis on the parameters' choices and the NN tuning process. Finally, the performances of the NNs are analyzed and some suggestions for future developments are reported to encourage further researches.

Abstract: italian version

Nell'ultimo decennio, con l'avvento dei veicoli elettrici ed ibridi, il comfort acustico all'interno dell'abitacolo è diventato uno dei fondamentali aspetti di progettazione dei veicoli autostradali con lo scopo di ridurre il livello del rumore prodotto e di migliorare la qualità del suono percepito. Il rumore all'interno dell'abitacolo di un veicolo è dovuto a tre cause principali: il motore, nel caso di autoveicoli con motori a combustione interna, il contatto delle ruote con il fondo stradale e l'aerodinamica del veicolo. Lo sviluppo di motori meno rumorosi e di veicoli in grado di viaggiare a velocità sempre più sostenute ha fatto emergere nuovi suoni che prima risultavano essere in qualche modo mascherati. Tra questi, il rumore dovuto dall'interazione del veicolo con l'aria circostante risulta essere importante già dai $100\ km/h$, diventando predominante con velocità superiori ai $130\ km/h$. I due fenomeni aeroacustici che maggiormente determinano *discomfort* acustico e fastidio durante la guida sono il *Leakage* and il *Whistle*. Il Leakage è un *broadband noise* ed è dovuto ad infiltrazioni di aria all'interno dell'abitacolo, quando questo non è perfettamente ermetico. Il Whistle è un *narrowband noise* ed è assimilabile ad un fischio dovuto a protuberanze dell'abitacolo, come ad esempio l'antenna della radio e gli specchietti retrovisori laterali. Gli attuali metodi per la rilevazione di questi fenomeni in veicoli a fine ciclo di produzione prevedono una prova su strada da parte di operatori che devono dare un giudizio soggettivo riguardo la presenza o meno di questi rumori ed eventualmente valutarne l'entità. Per rendere questo processo automatizzato, e di conseguenza più veloce e non soggetto ad errori dovuti a giudizi personali, si stanno sviluppando reti neurali in grado di rilevare tali rumori.

L'obiettivo di questa tesi, realizzata presso la *Siemens PLM Software,* è quello di sviluppare reti neurali che possano indicare la presenza o meno di Leakage e Whistle in veicoli a fine ciclo di produzione. In particolare, sono state adottate due reti neurali di tipo *Deep Neural Network* (DNN*)*. La prima, denominata *Leakage NN*, è stata allenata su un database con soli casi di Leakage e Non-Leakage ed ha il compito di rilevare il solo Leakage. La seconda, denominata *Whistle NN*, presenta il database con soli casi di Whistle e Non-Whistle ed ha lo scopo di individuare il Whistle. Le prestazioni di entrambe le reti neurali sono risultati più che promettenti, la Leakage

NN ha un'accuratezza sul database di test del 99.80%, mentre la Whistle NN ha raggiunto un'accuratezza sul database di test del 92.80%. Considerando che l'accuratezza di un operatore umano si aggira intorno al $60 - 70\%$, risulta del tutto evidente quali miglioramenti potranno apportare le reti neurali nei problemi NVH (Noise, Vibration, and Harshness) del prossimo futuro. Tuttavia, una delle principali limitazioni delle DNN, già note in letteratura e che sono state riscontrate anche in queste reti neurali, è la necessità di avere un grande database di allenamento composto da casi del tutto simili a quelli che dovranno essere predetti nella fase di test. Ad esempio, quando le due reti sono state allenate sui rumori prodotti da un Opel Vectra, non sono state in grado di rilevare il Leakage e il Whistle presente nel suono prodotto da una Ford Mondeo. Inoltre, la Leakage NN non è stata in grado di classificare i Whistle come casi di Non-Leakage, analogamente la Whistle NN non è stata in grado di catalogare i casi di Leakage come casi di Non-Whistle. Uno dei metodi a disposizione per migliorare le prestazioni delle reti neurali in situazioni differenti consiste nel generare database sempre più grandi e generalizzati, possibilmente con tanti modelli di auto e in diverse situazioni di funzionamento.

Acknowledgments

Vorrei ringraziare il mio relatore, il Prof. Paolo Castellini, per la sua disponibilità e per avermi permesso di svolgere il mio tirocinio presso la Siemens PLM Software in Belgio. Un'esperienza formativa inestimabile dal punto di vista professionale ma anche personale, nella quale ho avuto la possibilità di collaborare con persone con un alto livello di professionalità e competenza. Nonostante le difficoltà poste dal Covid-19, ho vissuto un'esperienza di sviluppo e ricerca in un contesto internazionale, acquisendo nuova maturità e consapevolezza.

Il mio secondo ringraziamento va al mio correlatore, il Dott. Claudio Colangeli, per avermi sempre guidata e consigliata durante tutto il mio tirocinio in Siemens. Inoltre, ha fatto sì che mi potessi ambientare velocemente in azienda, creando un rapporto professionale aperto al confronto ed al dialogo che si è poi trasformato in amicizia.

L'ultimo ringraziamento lo vorrei dedicare alla mia famiglia che mi ha sempre incoraggiata ad inseguire i miei sogni ed obiettivi con impegno e costanza, supportandomi anche in questa seconda esperienza estera

1 Introduction and Motivation

In recent times and, in particular, with the advent of electrification in automotive, vehicles interior acoustic comfort has become one of the main key factors in vehicle development process with the aim of reducing the level of interior noise and improving the sound quality as well. Since vehicles have continuously become quieter, the customer's sensitivity to acoustical comfort has increased due to two factors: on the one hand certain noise sources have become more audible due to reduced overall sound pressure levels and, on the other hand, due to the absence of very loud masking sounds, human hearing becomes more sensitive for any changes in the time and frequency domain (1).

The main sources of motor vehicle interior noise are engine noise, tire-road noise and aerodynamic noise (or wind noise). The drive train noise dominates at low speeds and high engine load, tire-road noise contributes the greatest part to the overall noise at low speeds and with a low engine load, while the aerodynamic noise of the vehicle becomes more significant at increasing speeds (2). With the growing popularity of electric and hybrid vehicles, where the engine noise has been drastically reduced, wind and tire-road noises have become predominant components of the vehicle interior noise. Moreover, as engine and road-tire noise has been minimized and as vehicles speed has increased over the last years, aerodynamic noise has become an important factor that influences the interior acoustic comfort. Leakages and whistling noises are the two main aerodynamic noises that determine discomfort and annoyance. Leakages in sealing systems and whistles can be represented by monopole sources and dipole sources respectively and, with a Mach number under 1, monopole sources are the most effective, followed by dipole sources (for more details see Chapter 2.1).

At the time of writing, leakages and whistles are detected and analyzed in a full vehicle End-of-Line using purely subjective methods which are error-prone and often time consuming. However, with the recent advances and transformative potential of Machine Learning (ML), including deep learning, in the field of acoustics (3), there have been numerous efforts to develop automated methods to identify and quantify NVH issues. The use of Neural Networks can drastically reduce vehicle evaluation time and sound engineers' efforts when compared with traditional subjective evaluation methods

and standard acoustic NVH (Noise, Vibration and Harshness) optimization procedures. A clear limitation of ML-based methods is that they are data-driven and thus require large amounts of data for testing and training (3). However, as the computing powers and cloud storages have been grown exponentially in the last decade, it is expected that data volumes will continue to increase.

This thesis describes the implementation of Neural Networks that can detect leakages and whistling noises in a full vehicle End-of-Line, with also a short description of data generation.

2 Theory

2.1 Causes for the generation of aerodynamic noise

The aerodynamic noise begin to influence the vehicle noise from a speed of approx. $100\ km/h$ and it dominates both exterior and interior noises from a speed of approx. $130\ km/h$ for passenger vehicles, and for small van-type trucks this threshold can be even lower (2).

Aerodynamic noise is essentially caused by three different mechanisms:

- pulsating volume flow through small openings that can be represented by monopole sources, such as leakages in sealing systems;
- impact pressure variations on hard surfaces that can be represented by a dipole source, such as whistling noises;
- turbulent shear stresses that creates quadrupole sources.

All these generation mechanisms have effects in the aeroacoustics of motor vehicles; however, the intensities of these three source types vary greatly (Figure 1).

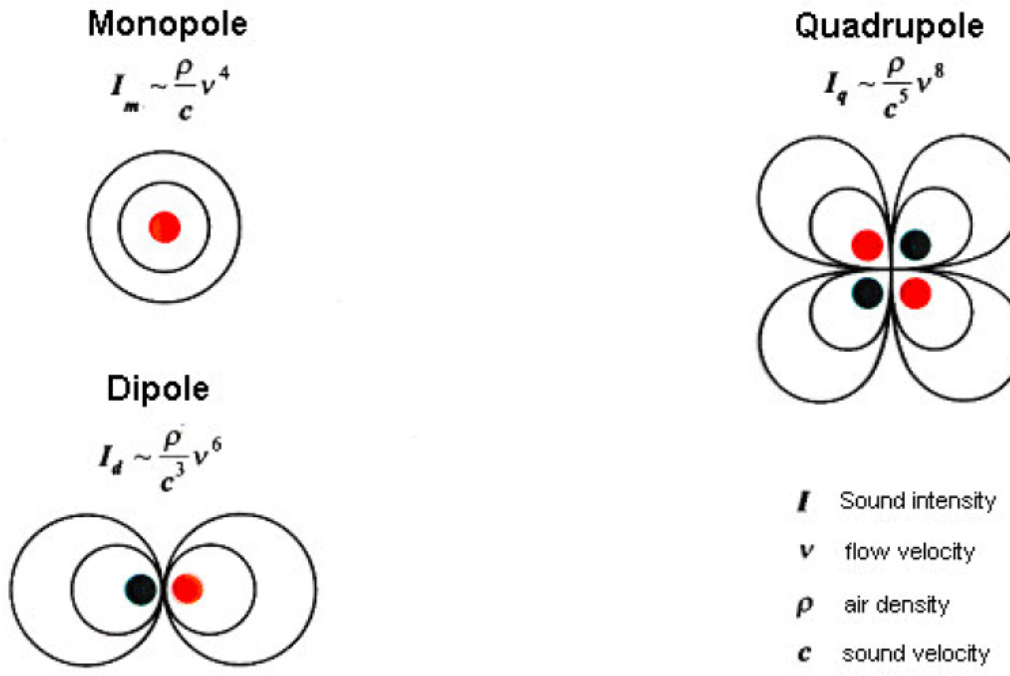


Figure 1 Schematic representation of the source types relevant in aeroacoustics. Adapted from Helfer (2005) (2).

With flow speed v, density ρ, sound velocity c and Mach number Ma, the intensity for a monopole source is:

$$I_m \sim \frac{\rho}{c} \cdot v^4 = \rho \cdot Ma \cdot v^3$$

And the intensity for a dipole source is:

$$I_d \sim \frac{\rho}{c^3} \cdot v^6 = \rho \cdot Ma^3 \cdot v^3$$

The sound power level of a monopole source is proportional to the 3rd power of the flow speed v, whereas the sound power level of a dipole source increases with the 6th power of speed.

Comparing the intensities at low flow speeds, with Mach numbers smaller than 1, the monopole source is the most effective, followed by the dipole source. The lowest emission is caused by the quadrupole sources, which can be neglected in most cases in vehicle aeroacoustics (2).

2.2 Leakage

The sealing system of a vehicle, particularly in the area of doors and side windows, can have a significant impact on the overall noise heard inside the cabin. Typical scenario is that the pressure difference between interior and exterior grows at higher speeds and negative pressures acting on the outside of the car could lift the doors out of their seals with a subsequent increase of leakage noise (Figure 2).

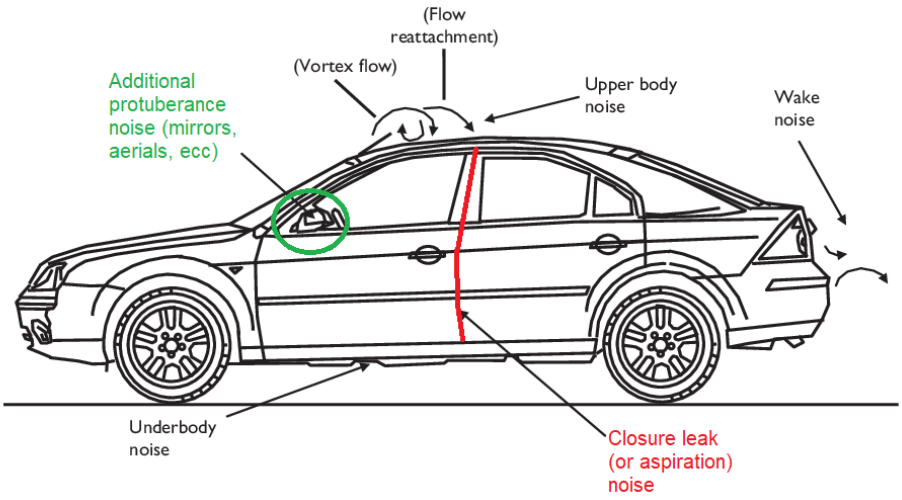


Figure 2 Schematic of the potential sources of aeroacoustic sounds in a vehicle environment.

As discussed in Chapter 2.1, the noise mechanism from leaks is that of a monopole, which has the highest intensity of all the aerodynamics noise sources. The leakage noise is also labelled as a *broadband noise* because its Power Spectral Density (PSD) has the power distributed over a wide section of frequencies; generally this energy per unit of time is located at higher frequencies (Figure 3).

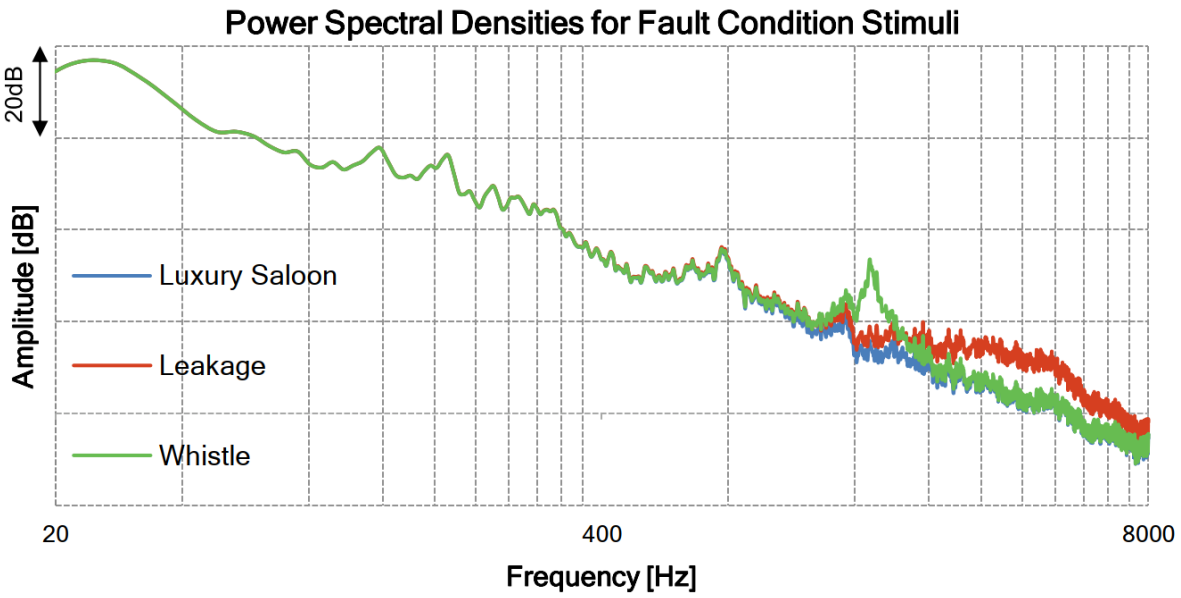


Figure 3 Power spectral densities for the original luxury saloon aeroacoustic signal and stimuli with a leakage and whistle added via filtering. The Leakage is a broadband noise thus the sound energy is distributed over a wide section of the audible range. The whistle is a narrowband noise thus the sound energy is distributed over a relatively small section of the audible range. Adapted from West B - 2019 (4).

2.3 Whistle

The add-on parts of a vehicle such as exterior mirrors and aerials can cause irritating whistling noises, due to the impact of a free or separated flow with hard surfaces (Figure 2). The formation of Von Karman vortices behind cylindrical bodies is an example of this type of acoustic source, where the vortex-induced pressure fluctuations create a dipole sound source on the body surface. The formation of Von Karman vortex street behind a radio aerial can be mitigated by increasing the angle of the aerial, wrapping the cylindrical profile with a helical strake or even removing the necessity for the device entirely (5). As discussed in Chapter 2.1, the noise mechanism from whistles is that of a dipole, which has the second highest intensity after leakage noises. The whistle is also labelled as a *narrowband noise (or tone)* because its PSD has the sound power distributed over a relatively small section of frequencies; generally this peak is located at higher frequencies (Figure 3).

2.3.1 St number

The frequencies resulting from the vehicle body and its add-on parts and details can be estimated with the equation

$$f = \frac{St \cdot v}{l}$$

where l is a characteristic measurement (e.g. height or width) of the individual vehicle component or detail and St is the Strouhal number. Generally, a Strouhal number of approx. 1 can be assumed for add-on parts, such as exterior mirrors. However, for parts with a cylindrical shape (e.g. aerials) it must be assumed as 0.2 (2).

2.4 Sound synthesis technique

As it will be discussed later, sound synthesis techniques will be adopted for generating digital twins of the acoustic behavior of produced cars. This will allow the generation of synthetic databases of car sounds originated from baseline recordings and properly modified in order to obtain a wide range of variations to be used for the purpose of training machine learning models. Sound synthesis approaches consist of three conceptual elements: (i) sound analysis and decomposition, (ii) sound editing and (iii) sound synthesis. Sound decomposition is the first step in creating a realistic virtual vehicle sound. Two decomposition philosophies are possible: top-down and bottom-up. The top-down approach consists in the application of algorithms for separation tonal and broadband noise components from sound recordings of full vehicles tested in operational conditions. The bottom-up approach, instead, leverages on techniques for combining spectral information about tonal and broadband noise contributions extracted for components, tested in benches or numerically simulated (Figure 4).

In this thesis, the top-down approach has been adopted. In this case, the first step is to decompose whole vehicle sounds into individual noise source contributions, such as tire-road, wind and engine (6). One of the main benefits of this approach is that a single measured sound can be decomposed into its component parts without the need for measuring road, wind and power-train noise separately (7). Moreover, this method allows to use sound modification tools to vary the contributions of each source and finally to re-synthesize the whole vehicle sound from the separate contributions for any driving condition.

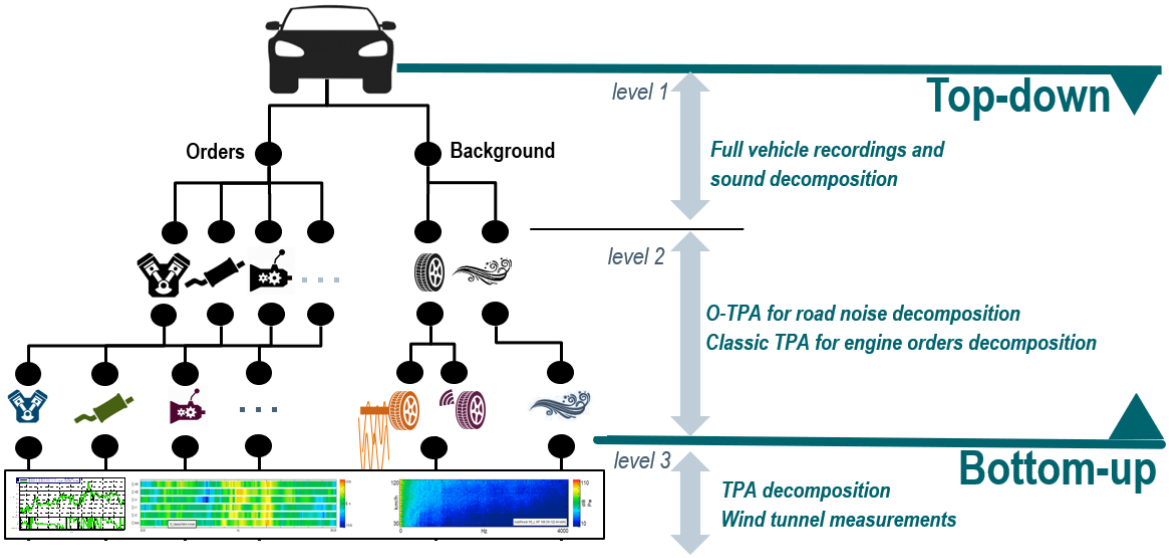


Figure 4 Vehicle sound decomposition and synthesis: Top-down and Bottom-up approaches. Top-down approach decomposes the full vehicle sound into different vehicle components. The Bottom-up method starts from the components data that are combined to form the full vehicle sound. Image courtesy of Siemens PLM Software.

The sound synthesis is based on a so-called Sound Quality Equivalent (SQE) model which is a collection of data that allows to synthesize a realistic vehicle sound in operational conditions based on a limited set of experimental recordings or numerical simulations. The vehicle sound is composed of tonal contributors or *orders*, whose frequency content is in a constant relation with the rotational speed of the engine, and by *broadband noise* or *background noise*, which are characterized by a broadband behavior in amplitude and a random nature in phase (Figure 5).

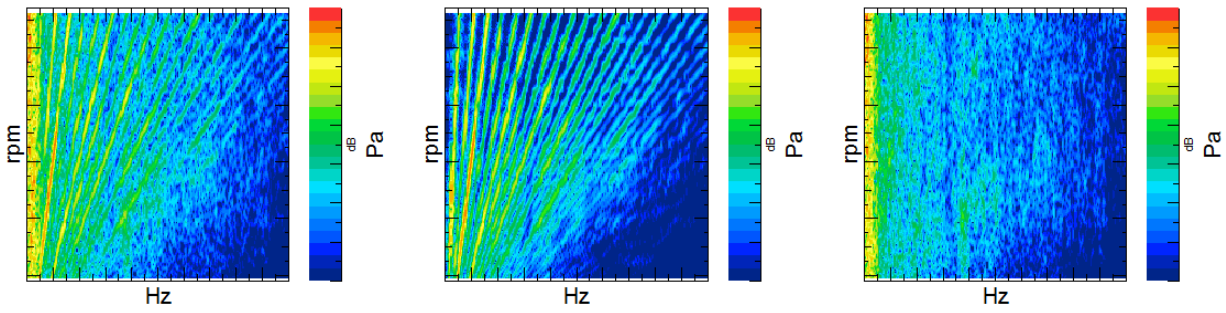


Figure 5 Vehicle sound decomposition into orders and broadband noise: total sound colormap (left), orders colormap (middle) and broadband noise colormap (right). Image courtesy of Siemens PLM Software.

2.4.1 Orders

The rotation of a component (e.g. crank shaft) emits vibrations and acoustic responses with a tonal behavior labelled as order (Figure 6). The order at which a rotating system is operating can be calculated from the rotational frequency f and the rotational speed rpm via the following equation:

$$O = f \cdot \frac{60}{rpm}$$

Frequency and orders are both a measure of events over an observation frame: the frequency is the number of events per unit of time and the order is the number of events per revolution (8).

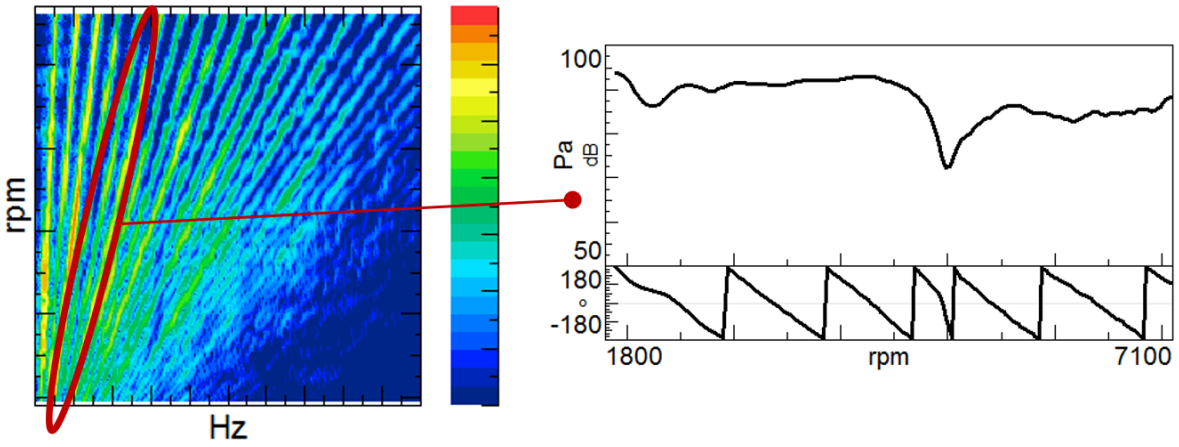


Figure 6 Example of order 3 of the Alfa Stelvio Quadrifoglio interior sound (WOT and 3rd gear). Image courtesy of Siemens PLM Software.

At the time of writing, there are three order tracking methods that can be used to identify the orders from a vehicle sound:

- Resampling based order tracking;
- Time Variant Discrete Fourier Transform (9);
- Vold-Kalman filter based order tracking (10).

2.4.2 Broadband noise

The wind noise and the tire-road noise have broadband nature. Their amplitude spectrum does not reveal tonal components and their phase has in general a random pattern (Figure 7).

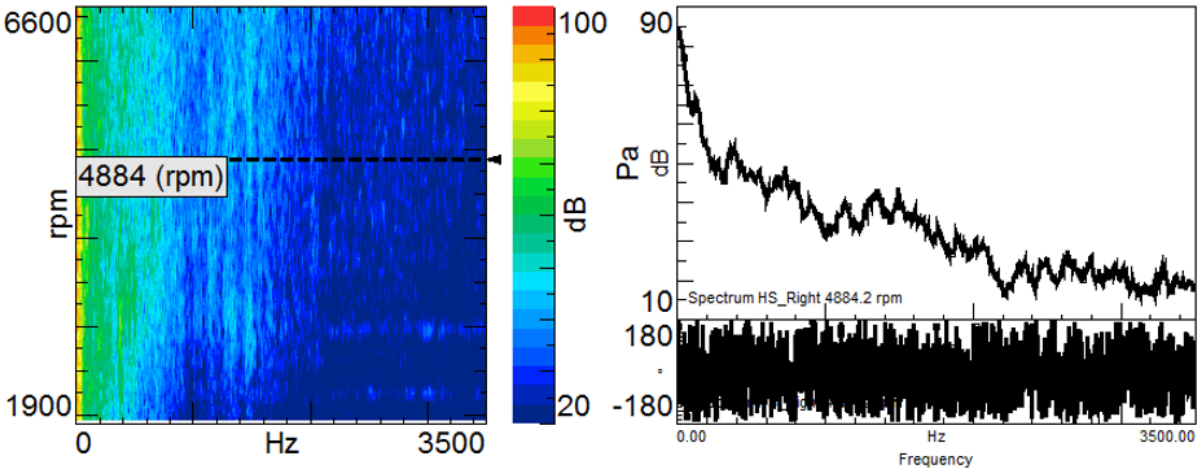


Figure 7 Example of broadband noise of the Alfa Stelvio Quadrifoglio interior sound (WOT and 3rd gear). Image courtesy of Siemens PLM Software.

At the time of writing, there are two broadband noise synthesis methods that can be used (11):

- Fractional octave synthesis;
- Narrow band synthesis.

2.5 MEL spectrogram

A Mel spectrogram is a spectrogram where the frequencies are converted to the Mel scale, which is a scale proposed by Stevens, Volkmann, and Newmann in 1937. The Mel scale is a unit of pitch such that equal distances in pitch sounded equally distant to the listener (Figure 8). Indeed, several studies have proved that humans do not perceive frequencies on a linear scale. The human hearing is better at detecting differences in lower frequencies than higher frequencies. For example, we can easily tell the difference between 500 and 1000 Hz, but we will hardly be able to tell a difference between 10,000 and 10,500 Hz, even though the distance between the two pairs are the same (12).

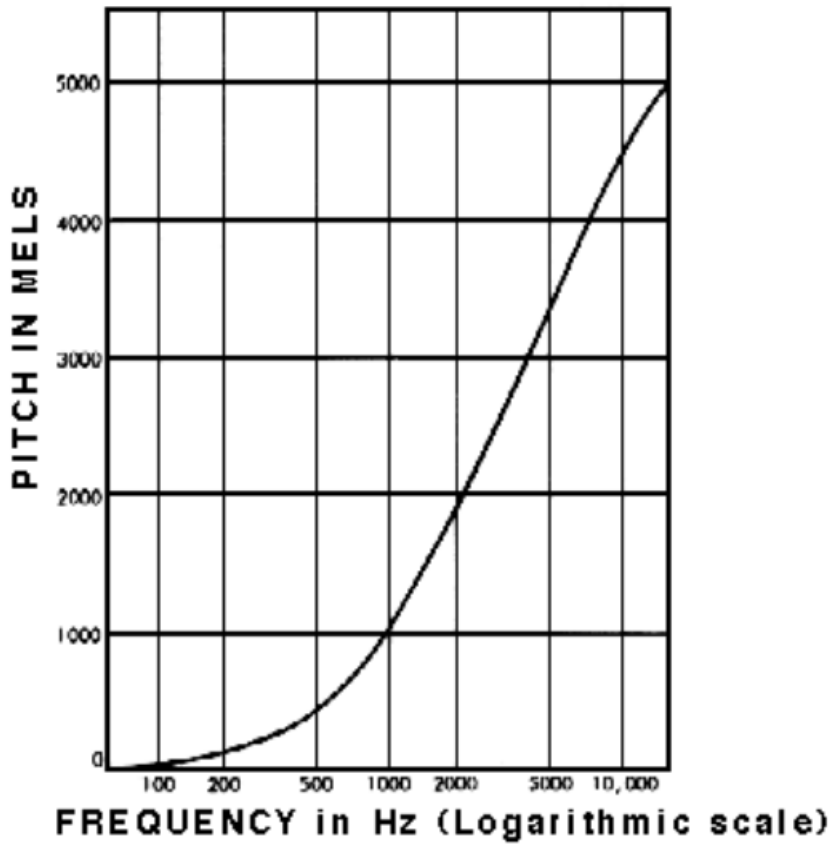


Figure 8 The Mel scale as a function of frequency. Adapted from Simon Fraser University.

2.6 Neural Networks

Machine Learning (ML) methods generally can be categorized as either supervised or unsupervised learning tasks. In supervised learning, the task is to learn a predictive mapping from inputs to outputs given labeled input and output pairs. Supervised learning is the most widely used ML category and includes familiar methods such as neural network (NN) models. In unsupervised learning, no labels are given, and the task is to discover interesting or useful structure within the data. This has many useful applications, which include data visualization, exploratory data analysis, anomaly detection, and feature learning (e.g. autoencoders) (3).

The purpose of feed forward NNs, also referred to as deep NNs (DNNs), is to approximate functions. These models are called feed-forward because information flows only from the *inputs* (or *features*) to the *outputs* (or *labels*), through the intermediate calculations. For more details see (3) and (13).

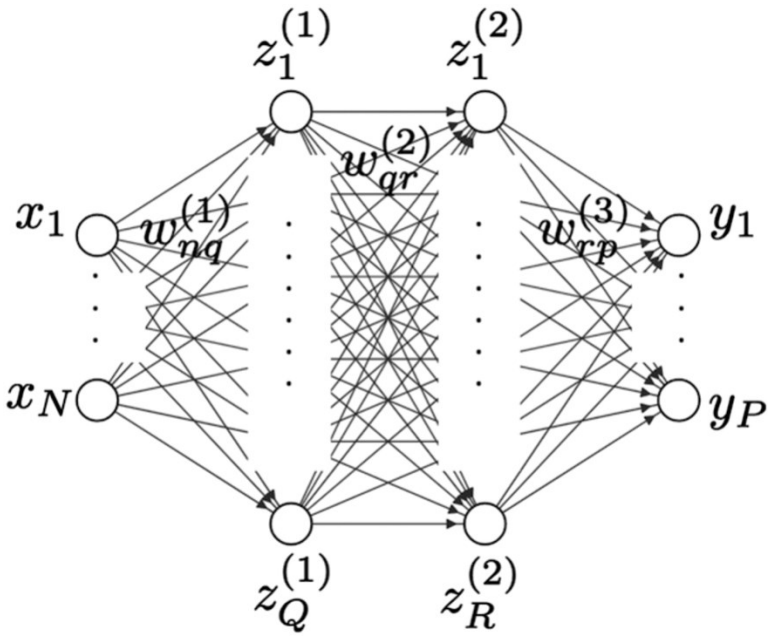


Figure 9 Example of three layers fully connected NN. Adapted from Michael J. Bianco - 2019 (3).

The number of layers in the network is the number of hidden layers plus one, which represent the output layer. The number of hidden units are the number of nodes inside each hidden layer.

The number of layers and the number of hidden units are two parameters (or hyperparameters) that affect the capacity of NNs, and they must be tuned to obtain the best performances. The activation functions are non-linear transformation of the inputs that are performed in the hidden and output layers; the ones that are chosen for this work are the Sigmoid and the Rectified Linear Units (ReLU). For more details see (3) and (13).

The model can be trained, tuned, and evaluated by dividing the data into three distinct sets: *training, validation*, and *test*. In this case the model is fitted on the training data, and its performance on the validation data is used to tune the hyperparameters (e.g. number of NN layers, number of hidden units, learning rate). Only after the hyperparameters are fully tuned on the training and validation data, the model performance is evaluated on the test data, which are unseen data that should never influence the model parameters. NNs perform best when their capacity is suited to the complexity of the data provided and the task, otherwise two situations can arise. If a high-capacity model is used for a low-complexity task, the model will *overfit* or learn the noise of the training set. In the opposite scenario, a low-capacity model trained on a high-complexity task will tend to *underfit* the data, or not learn enough details of the data. Both overfitting and underfitting degrade ML model *generalization*, which is the ability of the model to predict unseen data well. Underfitting and overfitting can be quantified using the *bias* and *variance* of the NN. The bias is the difference between the mean of our estimated targets and the true mean, and the *variance* is the expected squared deviation of the estimated targets around the estimated mean value (3).
A NN with high bias is underfitting the data and a NN with high variance is overfitting the data. Usually, the main things that help to reduce underfitting problems are to try bigger networks, such as more hidden units or more layers, or to train longer or even to try some optimization algorithms. The best way to solve a high variance problem is to get more data, otherwise the regularization could help to reduce overfitting (Figure 10). For more details see (13).

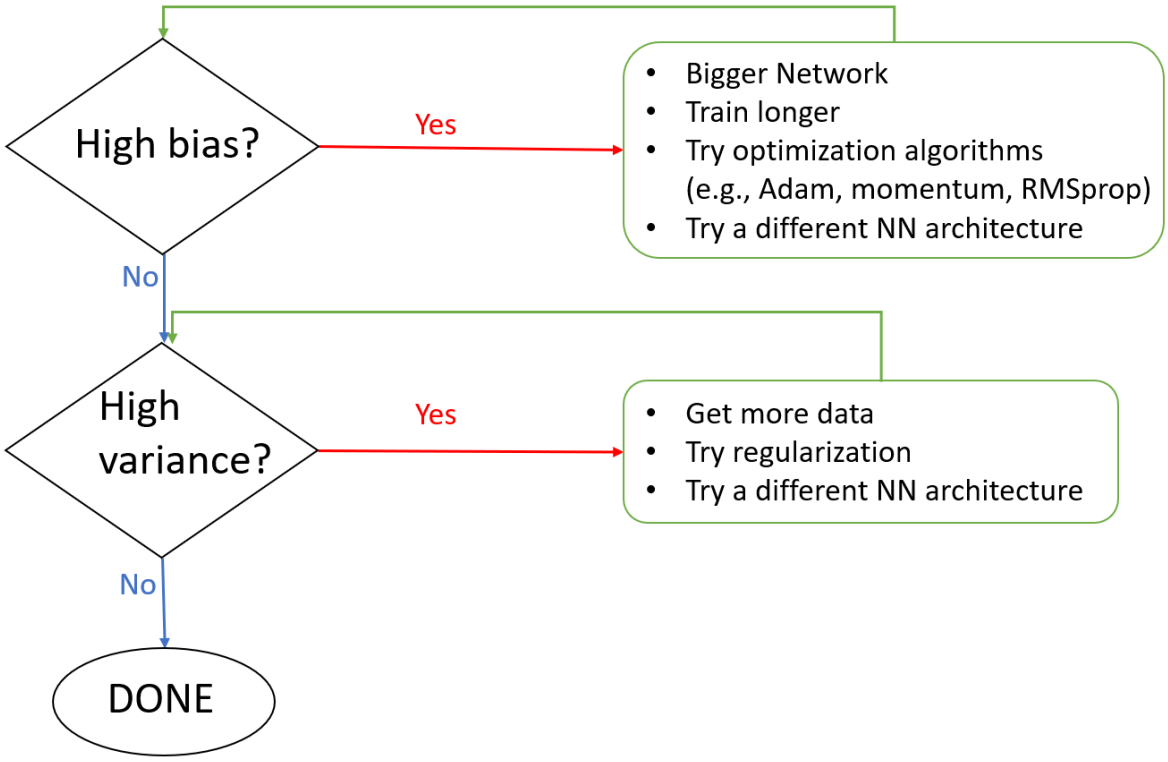


Figure 10 Flowchart of NN optimization process. Different actions could be attempted to avoid underfitting (or high bias) and overfitting (or high variance) problems. Note that the actions to reduce high variance has small effect on high bias, and vice versa.

2.6.1 Inputs normalization

To improve the NN performances, the training, the validation and the test sets were normalized with the following formula:

$$\frac{x-\mu}{\sigma} \qquad \mu = \frac{1}{n}\sum_{i=1}^{n} x^{(i)} \qquad \sigma = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(x^{(i)}-\mu)^2}$$

Where x is a vector with $32\ elements$, which corresponds to the MEL spectrogram frequencies of the input data (for more details see 2.5), μ is the mean and σ is the standard deviation. Note that to normalize the validation and test sets, one should use the same μ and σ calculated for the training set (for more details see (13)).

3 Executable digital twin of virtual vehicle for end-of-line database generation

In this work, vehicle sound synthesis techniques are adopted to generate executable digital twins of the acoustic behavior of studied vehicles in end-of-line. These executable digital twins allow to synthesize a realistic interior sound of end-of-line vehicle in operational conditions based on analytical and semi-empirical models. In this thesis, this approach was utilized to generate a big database of realistic interior sounds with whistles and leakages noises, required to train the Neural Networks.
With this method, it's possible to obtain a large number of realistic data from only one experimental recording (Figure 11).

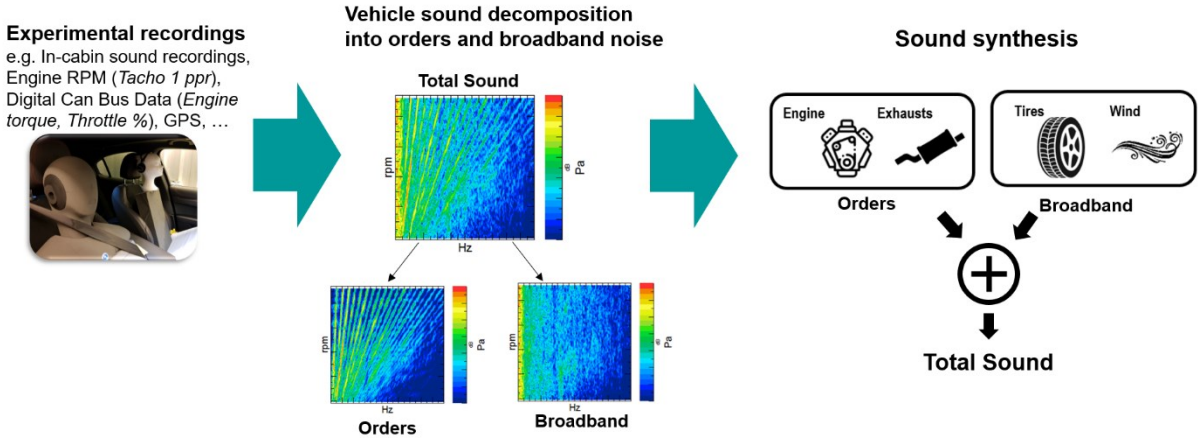


Figure 11 Sound synthesis techniques based on Vehicle Sound Simulator (VSS) technology. This method allows to use sound modification tools to vary the contributions of each source and finally to re-synthesize the whole vehicle sound from the separate contributions for any driving condition. Image courtesy of Siemens PLM Software.

First, the experimental sound was decomposed into orders and broadband noise with sound decomposition techniques (see Chapter 2.4), adopting the Vehicle Sound Simulator (VSS) software and technology (see Chapter 3.1). Next, the orders and broadband noise autopowers (AP) were exported from the VSS and imported in MATLAB, where the background noise was manipulated with the introduction of whistles and leakages (see Chapters 3.2, 3.3 and 3.4). Finally, the resulting vehicle sound has be synthesized and made available in time domain.

3.1 Vehicle Sound Simulator

With a Top-down decomposition in Simcenter Vehicle Sound Simulator software, the experimental noise was decomposed into background noise and orders (for more details see Chapter 2.4). The background noise is characterized by a broadband behavior in amplitude and a random nature in phase (Figure 12).

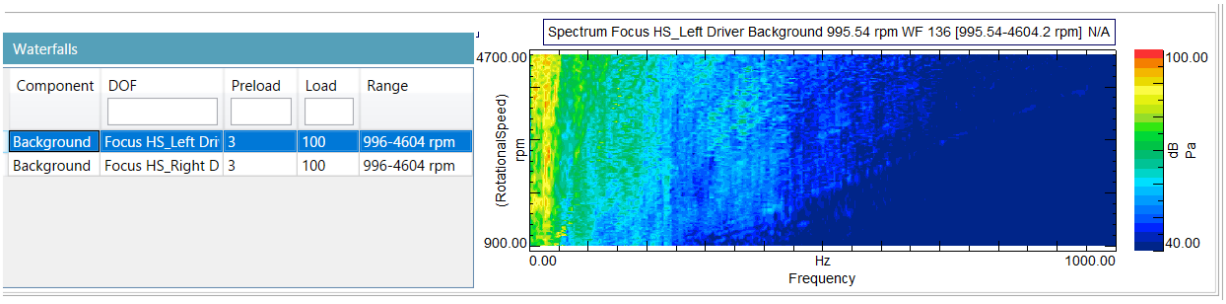


Figure 12 Background noise of a Ford Focus in WOT and 3rd Gear, decomposed by Simcenter Vehicle Sound Simulator. Image courtesy of Siemens PLM Software.

The tonal behavior of the order is noticeable at low frequencies, while at high frequencies the phase has some random features like broadband noises (Figure 13).

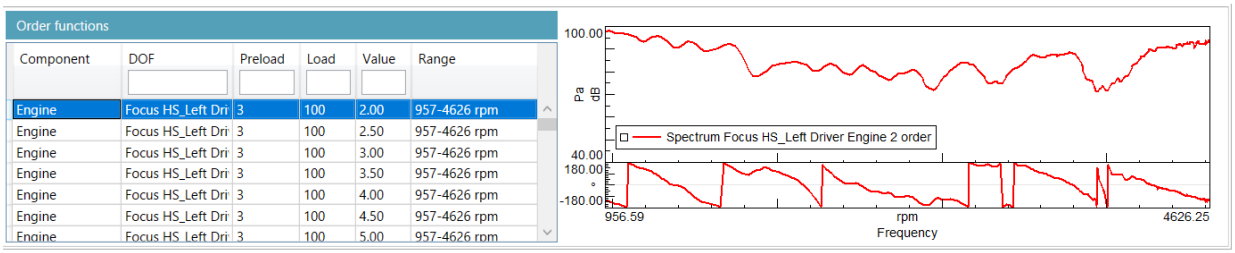


Figure 13 Amplitude and phase of the 2nd Order of a Ford Focus in WOT and 3rd Gear, decomposed by Simcenter Vehicle Sound Simulator. Image courtesy of Siemens PLM Software.

The SQE model obtained in VSS was imported into Testlab and then exported from Testlab as Excel files, ready to be used in MATLAB. In Figure 14, an example of a background AP that was exported from Testlab to MATLAB.

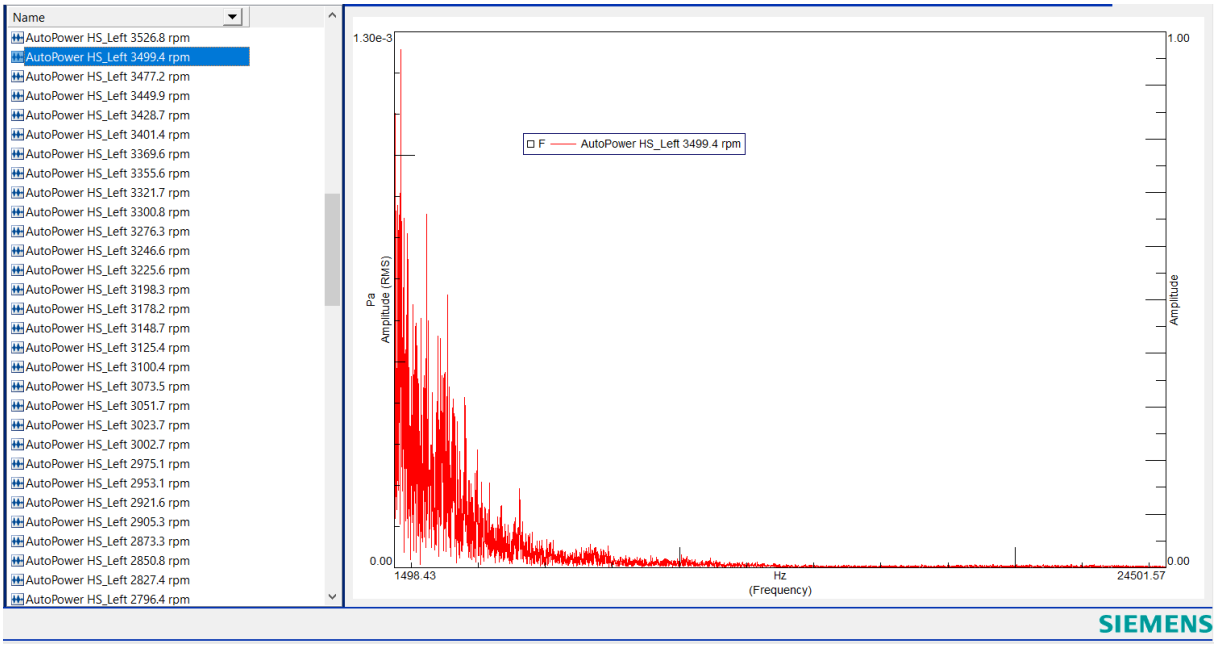
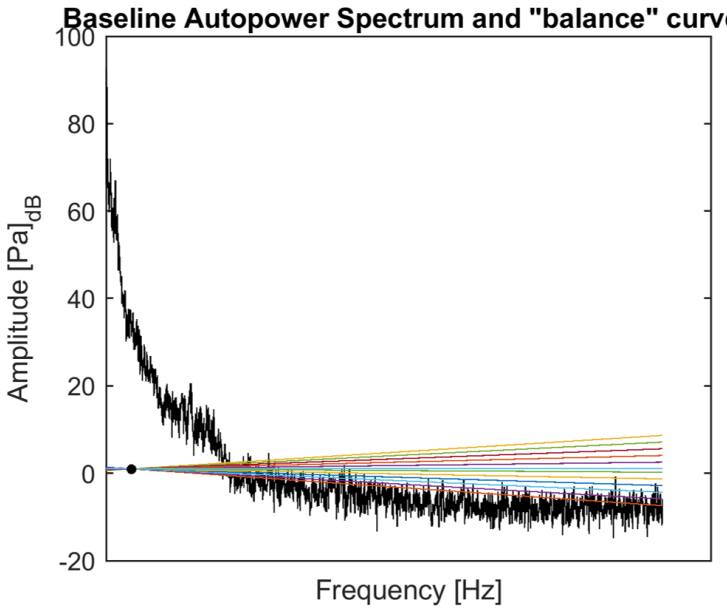


Figure 14 Background noise AP of a Ford Focus in WOT and 3rd Gear visualized in Simcenter Testlab. Image courtesy of Siemens PLM Software.

3.2 Non-Leakage and Non-Whistle cases

To obtain a better performance from the NN, the database must include not only cases with Leakages and Whistles (or *positive* cases) but also cases without Leakages and Whistles, called Non-Leakage and Non-Whistle cases or *negative* cases. Thus, to generate the Non-Leakage and Non-Whistle cases, the baseline background AP, imported from Testlab, was multiplied with a set of concurrent lines with random inclination angles lower than $0.025\,rad$. On top of that, random alteration within $6\,dB$ with respect to the given narrowband AP spectra has been applied. The combination of these two modelling approaches aim at mimicking the natural variations that likely happen when testing multiple vehicles of the same type and brand under the same nominal conditions. The AP baseline used in the study was obtained from the measured noise of an Opel Vectra (WOT and 3rd Gear) with a rotational speed of $3500\,rpm$. A set of $8500\,lines$ were generated with a pivot frequency and amplitude of $1000\,Hz$ and $1\,Pa$, respectively (Figure 15).



Data:

Opel Vectra

$Engine\ rotational\ speed = 3500\ rpm$

$Namber\ of\ generated\ cases = 8500\ cases$

$Pivot\ frequency = 1000\ Hz$

$Pivot\ amplitude = 1\ Pa$

$Maximum\ inclination\ \varphi = 0.025\ rad$

Figure 15 Opel Vectra Baseline Background noise Autopower Spectrum and "balance" curves, and the data used to generate the concurrent lines. The baseline background noise AP was imported from Testlab and then multiplied with the "balance" curves.

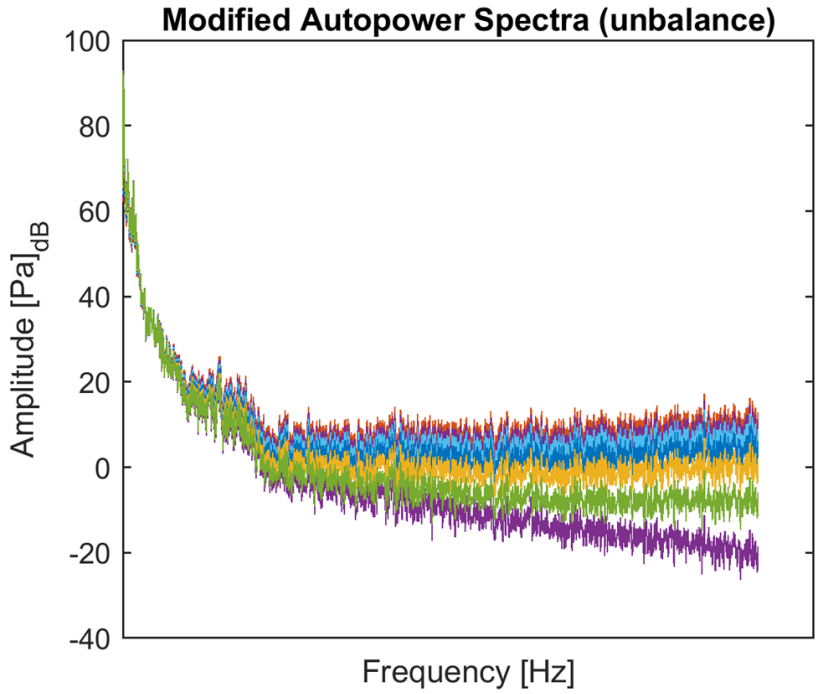


Figure 16 Autopower Spectra of the modified background noises, obtained from the multiplication of the baseline AP spectrum and the "balance" curves.

Therefore, 8500 Non-Whistle and Non-Leakage cases were generated from one experimental sound (Figure 16). The number of these negative cases was chosen related to the number of positive cases because balanced datasets are always preferable in Machine Learning (13).

3.3 Leakage cases

As discussed in Chapter 2.2, the leakage PSD has the power distributed over a wide section of frequencies (Figure 3). To replicate the leakage PSD shape, a new function was interpolated from the vertices of a trapezium (Figure 17 and Figure 18). Such a trapezoidal-shaped feature has been added to the baseline broadband noise spectra of the studied car in order to obtain synthesized sounds including leakage phenomena. A large range of leakage sounds have been generated. Taking as reference Figure 17, where the modelling parameters are depicted, the leakage database includes cases for which the ΔF parameter spans all values within the range between $1100 Hz$ and $4000 Hz$ with a step of $100\ Hz$. The severity of the leakage phenomenon is modelled through the amplitude of the B and C vertices, whose values range between $0.0006\ Pa$ and $0.001\ Pa$ (Figure 18).

The ranges used for amplitude and frequency were determined after an analysis of real noises with leakages measured in different types of car.

The ellipses' parameters were selected to preserve the trapezoidal shape after interpolation. For example, the gap Δf must be higher than the ellipse's semi-minor axis ${r_x}/{2}$ (Figure 18). The size and the characteristics of the database have been chosen to avoid overfitting and underfitting in the NN (for more details see Chapters 2.6 and 3.6).

Data:

$Point\ D\ range\ of\ frequency$

$D_{min} = 1100\ Hz$

$D_{max} = 4000\ Hz$

$D_{step} = 100\ Hz$

$\Delta F\ range\ of\ frequency$

$E_{min} = 2000\ Hz$

$E_{max} = 4000\ Hz$

$E_{step} = 100\ Hz$

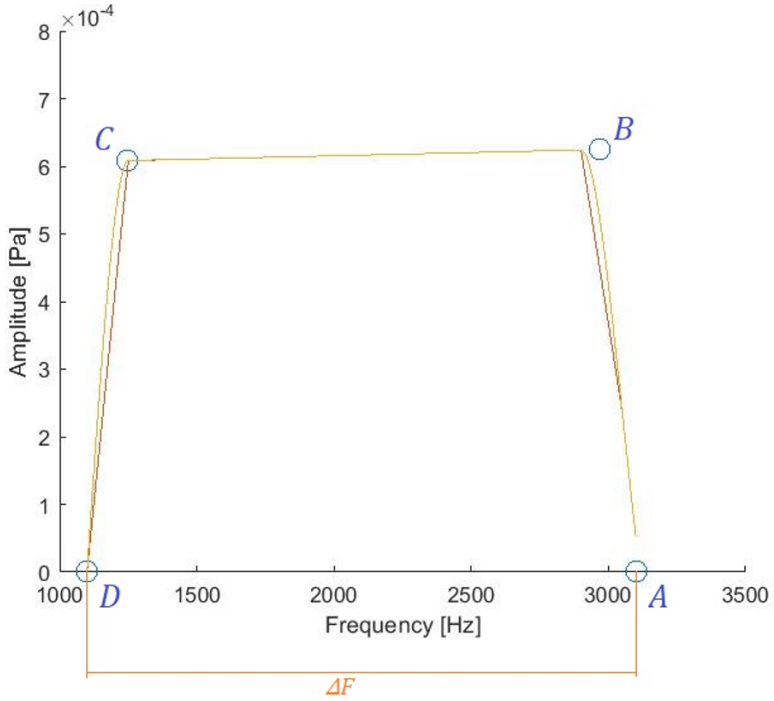


Figure 17 Leakage function interpolated from the vertices of a trapezium: the vertices change at every iteration of a for loop: D and A were selected at every step in the range of frequencies here reported, while B and C were randomly chosen from the two ellipses represented in Figure 18. The ranges used for amplitude and frequency were chosen after an analysis of real noises with leakage measured in different types of car.

Data:

$Height\ H\ range\ of\ frequency$

$(typical\ value = 0.000632\ Pa)$

$H_{min} = 0.0006\ Pa$

$H_{max} = 0.001\ Pa$

$H_{step} = 0.0001\ Pa$

$Ellipse\ parameters$

$\Delta f = 150\ Hz$

$r_x = 100\ Hz$

$r_y = 0.0002\ Pa$

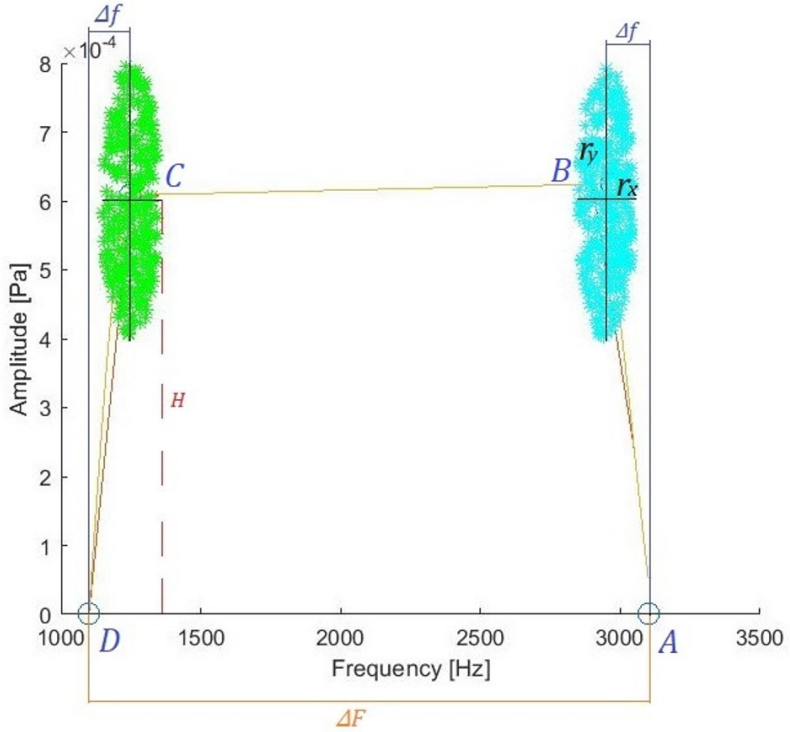


Figure 18 Leakage function interpolated from the vertices of a trapezium. The vertices change at every iteration of a for loop: B and C were randomly chosen from the two ellipses here reported, while the vertices D and A were selected at every step in the range of frequencies reported in Figure 17. The ranges used for amplitude and frequency were chosen after an analysis of real noises with leakage measured in different types of car.

Initially, the interpolation of the leakage function was realized using the MATLAB function *polyval*, but the cubic function introduced minimum points when there was a small gap between the vertices B and C. Therefore, another interpolation method was adopted to avoid the presence of minimum points in the leakage function. Firstly, a linear interpolation was used to increase the number of trapezium's points. Then, a second interpolation with the MATLAB function *interp1* was used to generate a cubic function. This method resulted to be not only the most stable but also the most efficient in term of time (Figure 19). Next, some noise was added to the interpolated function to obtain a realistic leakage noise. Finally, the leakage noise spectrum was added to the baseline background spectrum, which was previously modified using the method with the "balance" curves (for more details see Chapter 3.2).

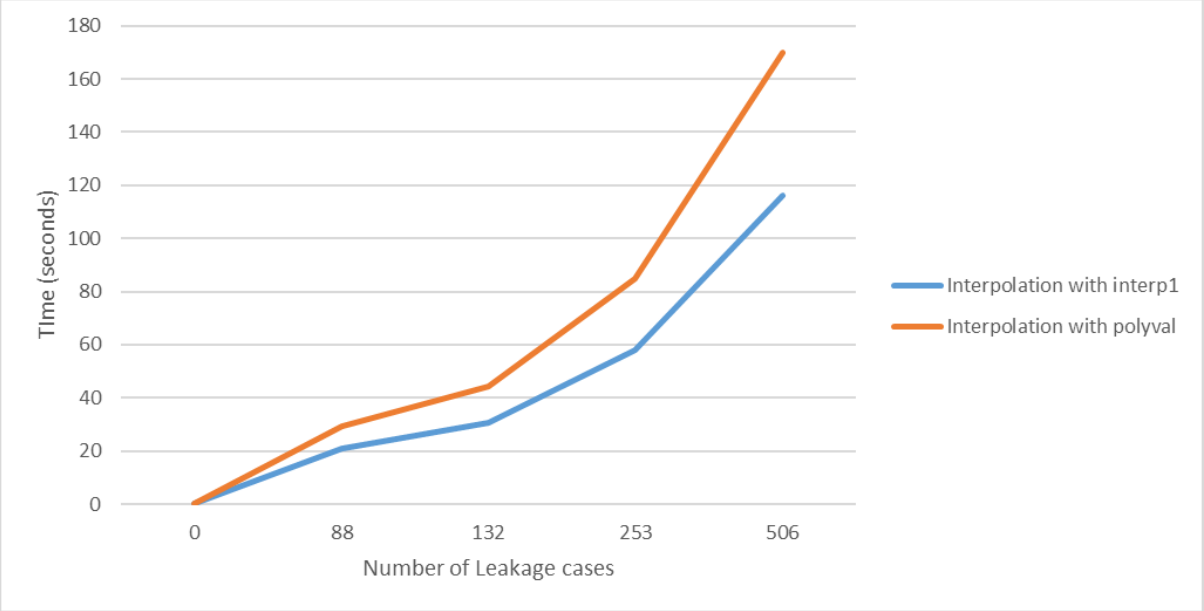


Figure 19 Comparison between "interp1" and "polyval" MATLAB interpolation methods. The interpolation method using "interp1" in MATLAB has the best performances not only in terms of time but also in term of stability. Indeed, the "polyval" function sometimes generates a leakage function with minimum points, which must be avoided.

3.4 Whistle cases

As discussed in Chapter 2.3, the whistle PSD has the power distributed over a relatively small section of frequencies (Figure 3). The peak of the whistle was obtained with the Lorentzian (or Cauchy) function and then added to the baseline broadband noise spectra of the studied car in order to obtain synthesized sounds including whistling phenomena. A large range of whistle sounds have been generated. Taking as reference Figure 20, where the modelling parameters are depicted, the whistle database includes cases for which the frequency of the peak spans all values within the range between $244\ Hz$ and $4444\ Hz$ with a step of $150\ Hz$. The severity of the whistle phenomenon is modelled through the amplitude of the peak, whose values range between $0.001\ Pa$ and $0.007\ Pa$. The shape of the whistle is chosen with the width at half amplitude between $10\ Hz$ and $160 Hz$ (Figure 20).

The ranges used for amplitude and width were determined after an analysis of real noises with whistles measured in different types of car. The typical peak's frequencies were determined by the St number (for more details see Chapters 2.3.1 and 3.4.1). The size and the characteristics of the database have been chosen to avoid overfitting and underfitting in the NN (for more details see Chapters 2.6 and 3.6).

Some noise was added to the Lorentzian function to obtain a realistic whistle noise. Finally, the whistle noise spectrum was added to the baseline background spectrum, which was previously modified using the method with the "balance" curves (for more details see Chapter 3.2).

Lorentzian function data:
Width at half amplitude
$A_{min} = 10\ Hz$
$A_{max} = 160\ Hz$
$A_{step} = 10\ Hz$

The peak amplitude
(typical values $0.001 \div 0.0063\ Pa$*)*
$B_{min} = 0.001\ Pa$
$B_{max} = 0.007\ Pa$
$B_{step} = 0.001\ Pa$

The peak frequency
(typical values $288 \div 4444\ Hz$*)*
$C_{min} = 244\ Hz$
$C_{max} = 4444\ Hz$
$C_{step} = 150\ Hz$

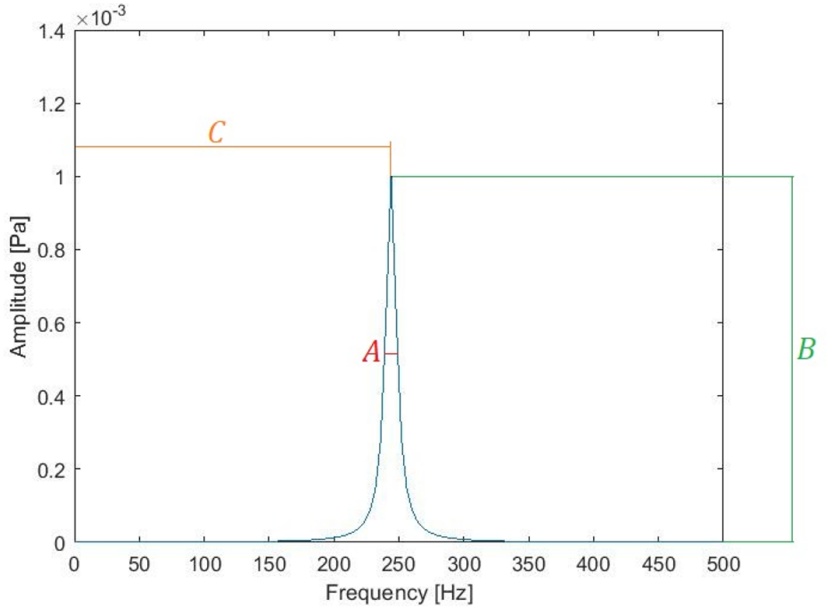


Figure 20 Whistle's peak obtained with the Lorentzian function and its parameters ranges. The typical values of the width and the peak's amplitude were chosen after an analysis of real noises with whistles measured in different types of car. The frequency typical values were calculated with the St number (for more details see Chapters 2.3.1 and 3.4.1).

3.4.1 Typical Whistle frequencies calculated with St

The typical frequencies of Whistles resulting from the vehicle's add-on parts and details can be estimated with the Strouhal number (see Chapter 2.3.1).

In particular, the diameter of add-on parts with cylindrical shapes ($St = 0.2$) was considered between $5\ mm$ to $2.5\ cm$ with a wind speed between $130\ km/h$ and $160\ km/h$. The values of the characteristic measurement (e.g. height or width) for any other vehicle component were varied between $10\ cm$ to $30\ cm$. The maximum and minimum frequencies resulted from these hypotheses were $4444\ Hz$ and $288\ Hz$, respectively.

3.5 Machine learning database composition

As the sounds of Whistle, Leakage, Non-Whistle and Non-Leakage cases were synthesized in the time domain, a *feature extraction* was performed in order to obtain the NN inputs features. There were many methods available in signal processing but, given the broadband nature of the phenomena studied in this thesis, two valid alternatives have been evaluated: the first one was the AP spectra and the second one was the Mel spectrograms. The AP spectra are more detailed because contain narrow band information, but they are also quite heavy in terms of memory allocation. For example, $11608\ cases$ of Whistle occupies $57.2\ GB$ of memory as AP spectra, and only $0.649\ GB$ as Mel spectrograms (Figure 21). Moreover, the NNs complexity (e.g. number of hidden layers and number of hidden units) have been decreased because a Mel spectrogram has only 32 univocal frequencies, while for example a PSD spectrum has usually a higher number of frequencies that depends on the level of discretization chosen for the frequency domain. As shown in Figure 22, despite the only 32 input features, the Whistle and Leakage phenomena are still recognizable. Therefore, the MEL spectrograms coefficients were chosen as NN input features.

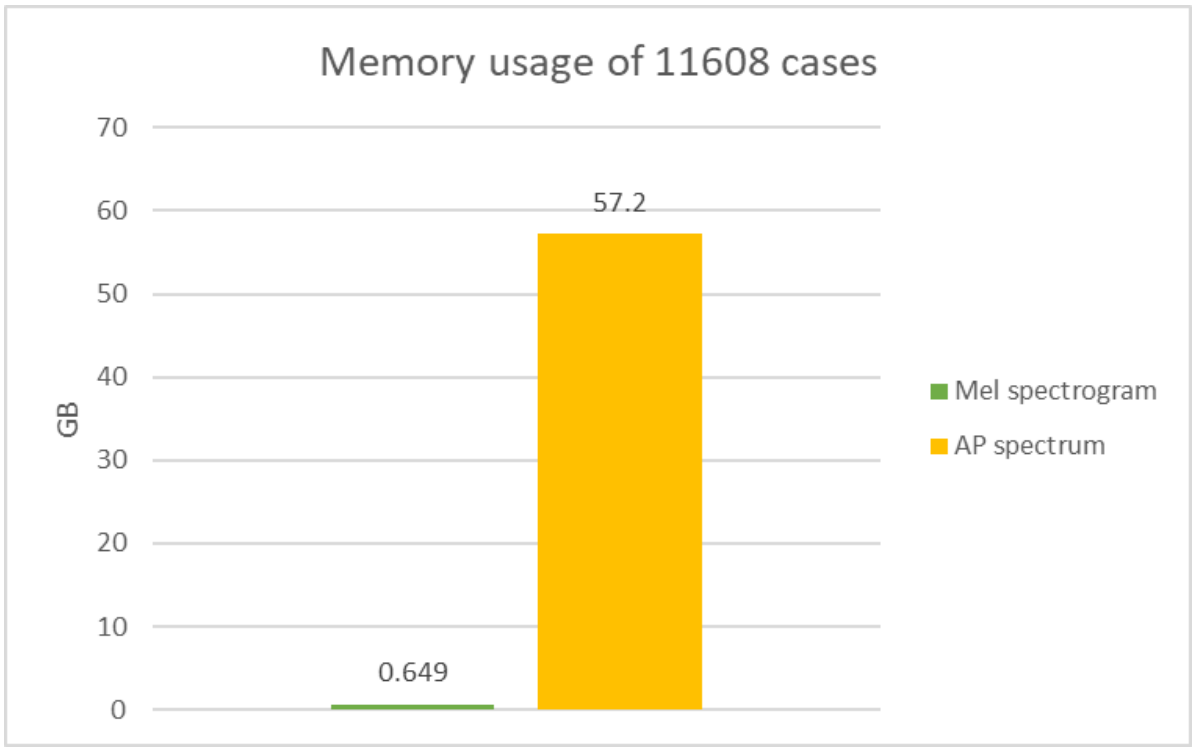


Figure 21 Comparison between Mel spectrogram and AP spectrum memory usage in GB. The Mel spectrogram not only permits to reduce drastically the memory usage but also to improve the NN performances.

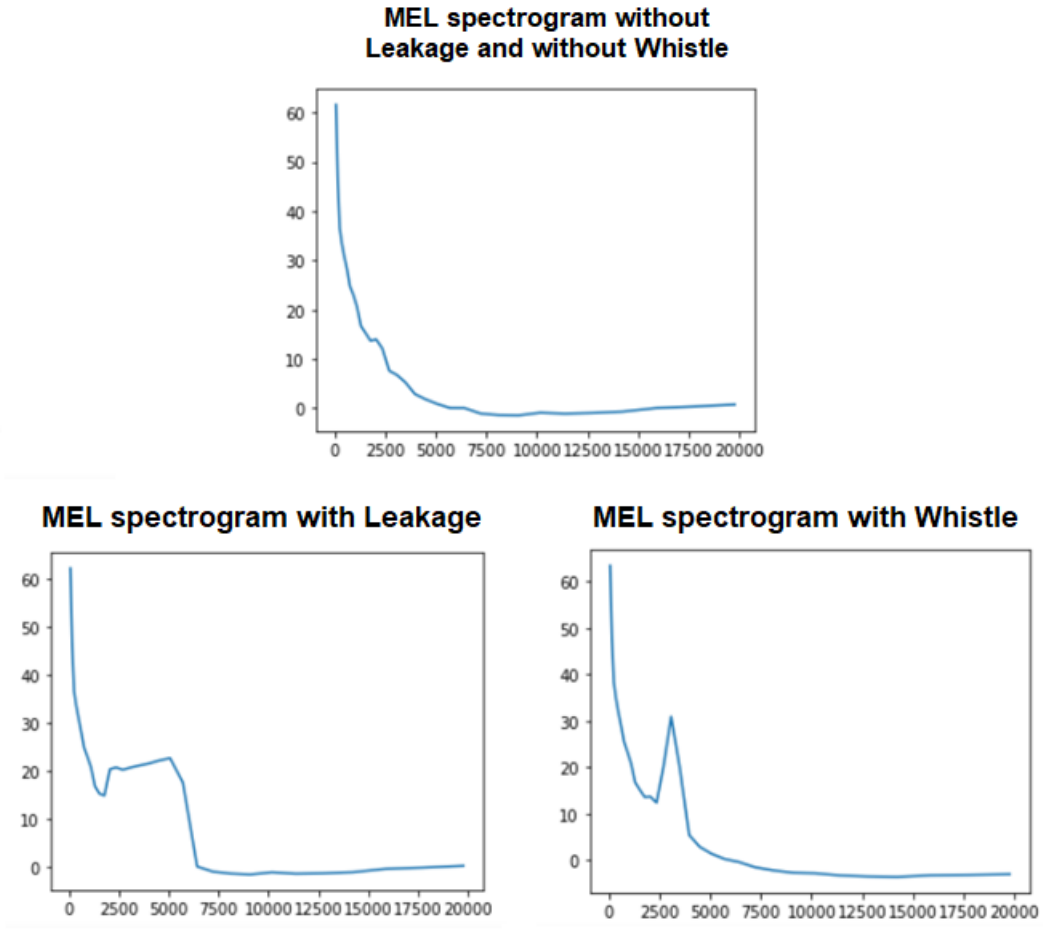


Figure 22 Non-Leakage and Non-Whistle Mel spectrum, Leakage Mel spectrum, Whistle Mel spectrum.

The model data were divided into three distinct sets: *training, validation*, and *test*. The model was fitted on the training data, and its performance on the validation data was used to tune the hyperparameters. Only after the hyperparameters were fully tuned on the training and validation data, the model performance was evaluated on the test data, which are unseen data that should never influence the model parameters (for more detail see Chapters 5.2 and 5.3). In Machine Learning balanced datasets are always preferable (13), however sometimes it is possible to obtain an unbalanced dataset that achieve superior NN performances.

3.6 Leakage Database

After an iterative process, the database that achieved better performances and more balanced datasets than any other has 12535 Leakage cases (59%) and 8750 Non-Leakage cases (41%) (Figure 23). Indeed, with 8750 Leakage cases and 8750 Non-Leakage cases the dataset would be balanced but the NN performances would be worse than the chosen case (for more details see Chapters 5.1.3).

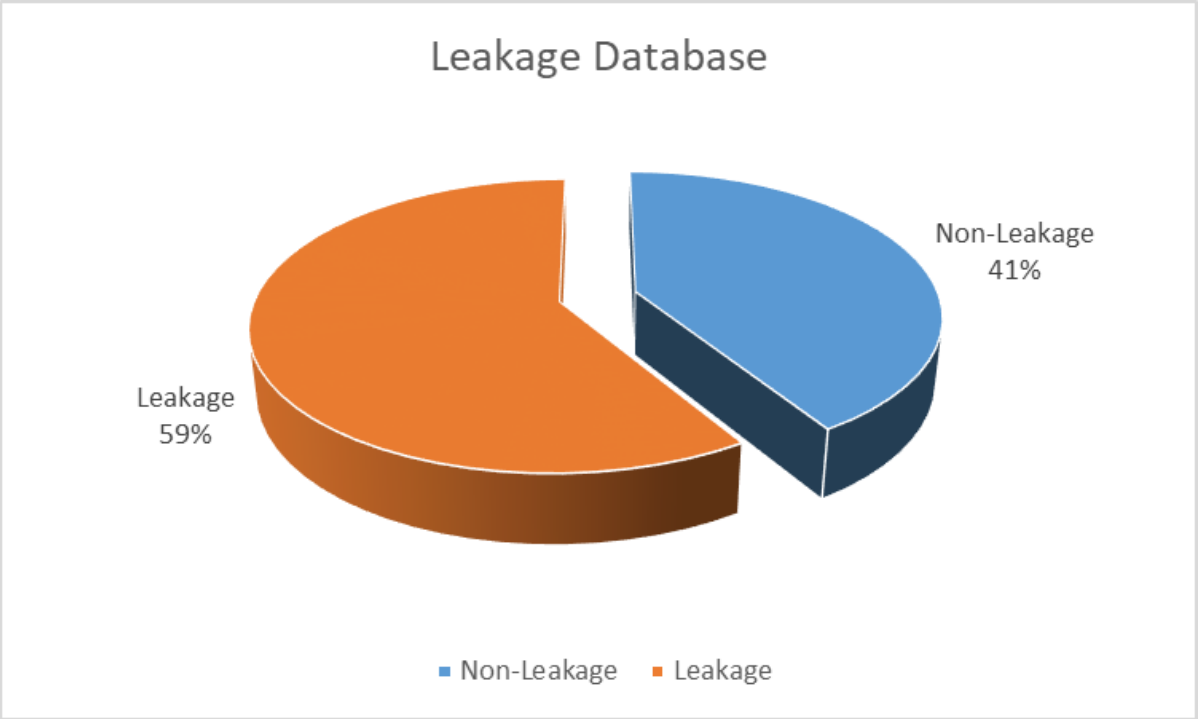


Figure 23 Composition of Leakage database chosen after an iterative process with the aim to achieve the best NN performances.

3.6.1 Leakage Train, Validation and Test sets

In Figure 24 is reported the partitioning of the Leakage Database that gives the best NN performances (for more details see Chapters 5.1 and 5.2). The Leakage Database has in total 21287 cases, whose 20285 are training cases (95%), 500 are Validation cases (2%) and 502 are Test cases (3%).

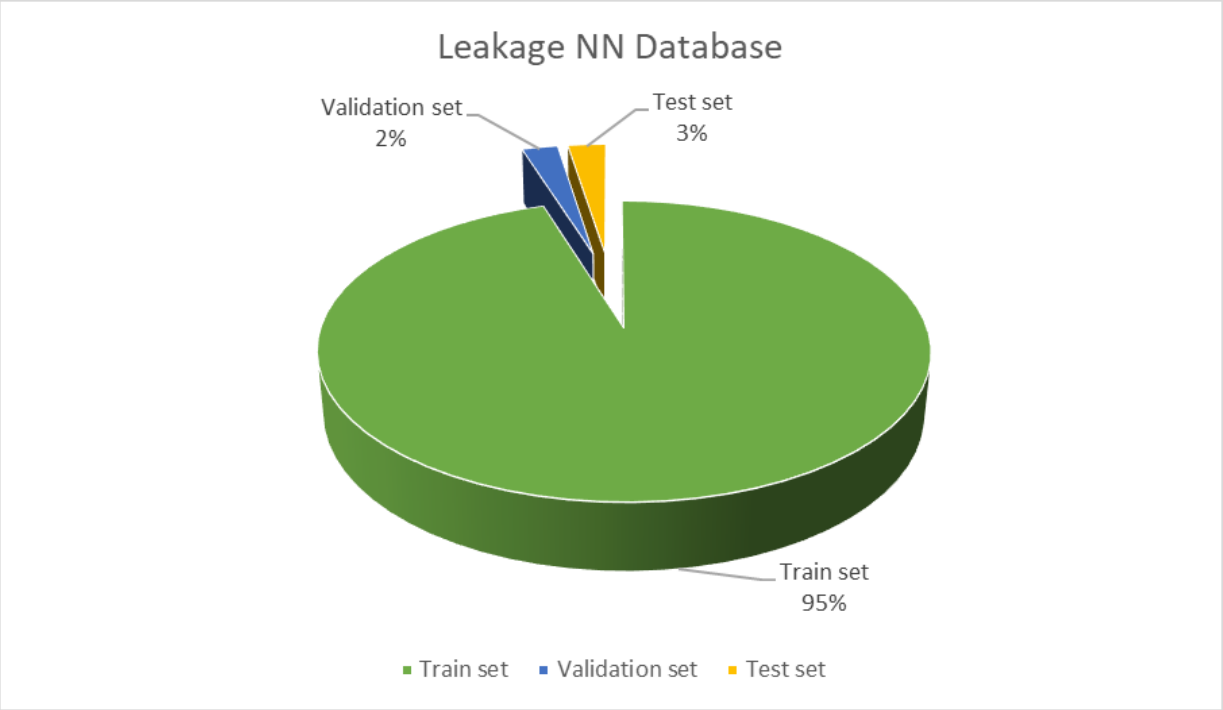


Figure 24 Leakage Database divided into Train, Validation and Test Sets. The number of Training, Validation and Test sets are the solution that gives the best NN performances, for more details see Chapter 5.1.

3.7 Whistle Database

As seen for the Leakage Database, the Whistle Database was chosen after an iterative process with the aim to obtain the best performances and the most balanced datasets. The database that achieved better performances and more balanced datasets than any other has 12960 Whistle cases (59%) and 9000 Non-Whistle cases (41%) (Figure 25). Indeed, with 9000 Whistle cases and 9000 Non-Whistle cases the dataset would be balanced but the NN performances would be worse than the chosen case (for more details see Chapter 5.3.3).

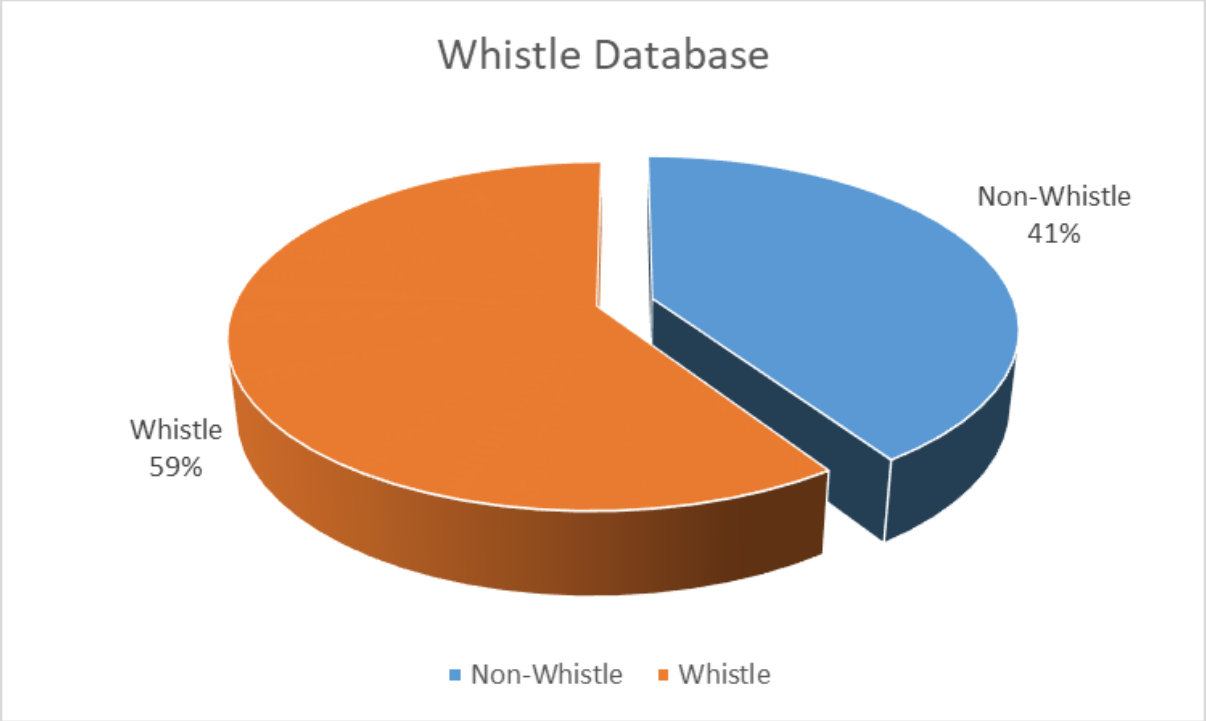


Figure 25 Composition of Whistle database chosen after an iterative process with the aim to achieve the best NN performances.

3.7.1 Whistle Train, Validation and Test sets

In Figure 26 is reported the partitioning of the Whistle Database that gives the best NN performances (for more details see Chapters 5.3 and 5.4). The Whistle Database has in total 21960 cases, whose 19932 are training cases (91%), 1000 are Validation cases (4%) and 1028 are Test cases (5%).

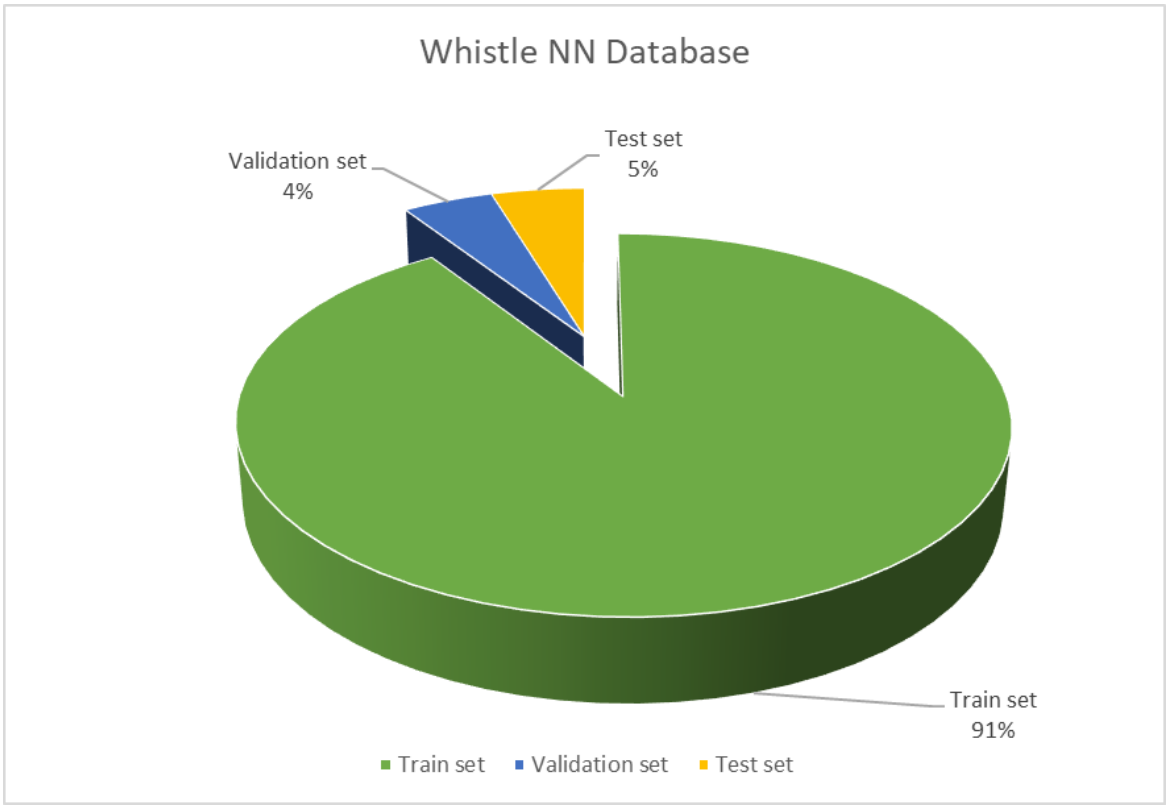


Figure 26 Whistle Database divided into Train, Validation and Test Sets. The number of Training, Validation and Test sets are the solution that gives the best NN performances, for more details see Chapter 5.3.

4 Wind noise classifier definition

To classify and detect input features, the supervised learning is the most suitable method because its task is to learn a predictive mapping from inputs to outputs, given labeled input and output pairs. Supervised learning can be applied in many different problems, thus many types of NN have been developed. Among them, the most common neural networks used in sound classification are the Deep Neural Networks and the Convolutional Neural Networks (CNNs). The main difference between DNNs and CNNs is the type of inputs: images for the CNN and feature vectors for the DNN. As the Mel spectra can be considered as vectors with 32 elements, the DNN technology results to be the most appropriate (Figure 27).

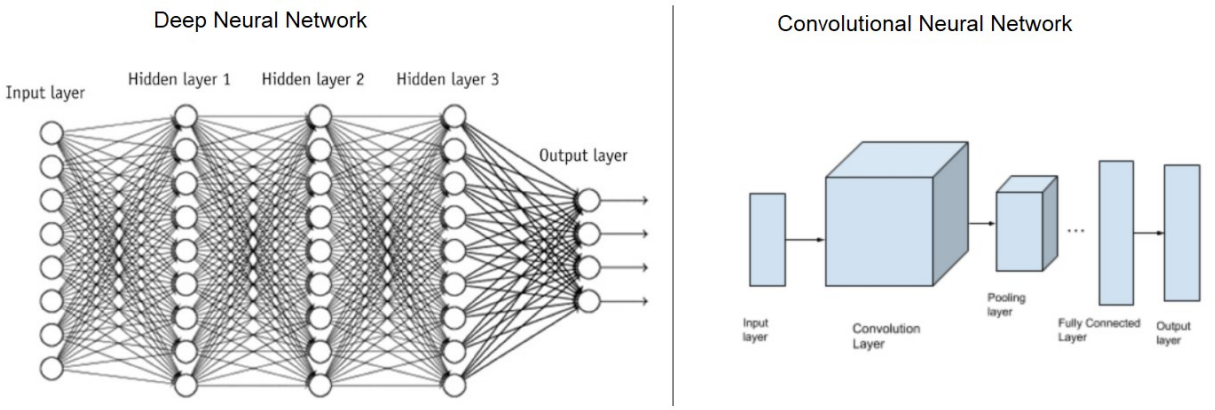


Figure 27 Deep Neural Network and Convolutional Neural Network structures. The input layer of the DNN is composed by feature vectors, while the input layer of the CNN is an image.

Two valid approaches were available: the first one was to use two different DNN with two outputs, and the second one was with one DNN with three outputs. The first procedure provides one DNN for the classification of leakages (Leakage and Non-Leakage) and the other one for whistles (Whistle and Non-Whistle). The second strategy allows to predict Whistle, Leakage, and Non-Whistle_Non-Leakage cases with only one intricate network. At the time of writing there are only few studies about DNN in sound classification and no papers about wind noise classifiers in automotive applications. Thus, to avoid any mistakes due to the few information available and the high difficulty of the network, the method with two different DNN was adopted in this thesis.

4.1 Deep Neural Network

The Mel spectra were exported from MATLAB as H5 files, which are data files saved in the Hierarchical Data Format (HDF), and then imported in Python, where the DNN was modeled. Taking as reference Figure 28, where a simplified representation of the Whistle DNN is reported, the 32 Mel frequencies corresponds to the input features of the NN. The DNN is composed by $(L-1)$ *Linear-ReLU functions* and a *Linear-Sigmoid function* for the output layer L. When the sigmoid result is higher than 0.5 means that the probability that the input has a whistle is higher than the probability that it's a non-Whistle case.

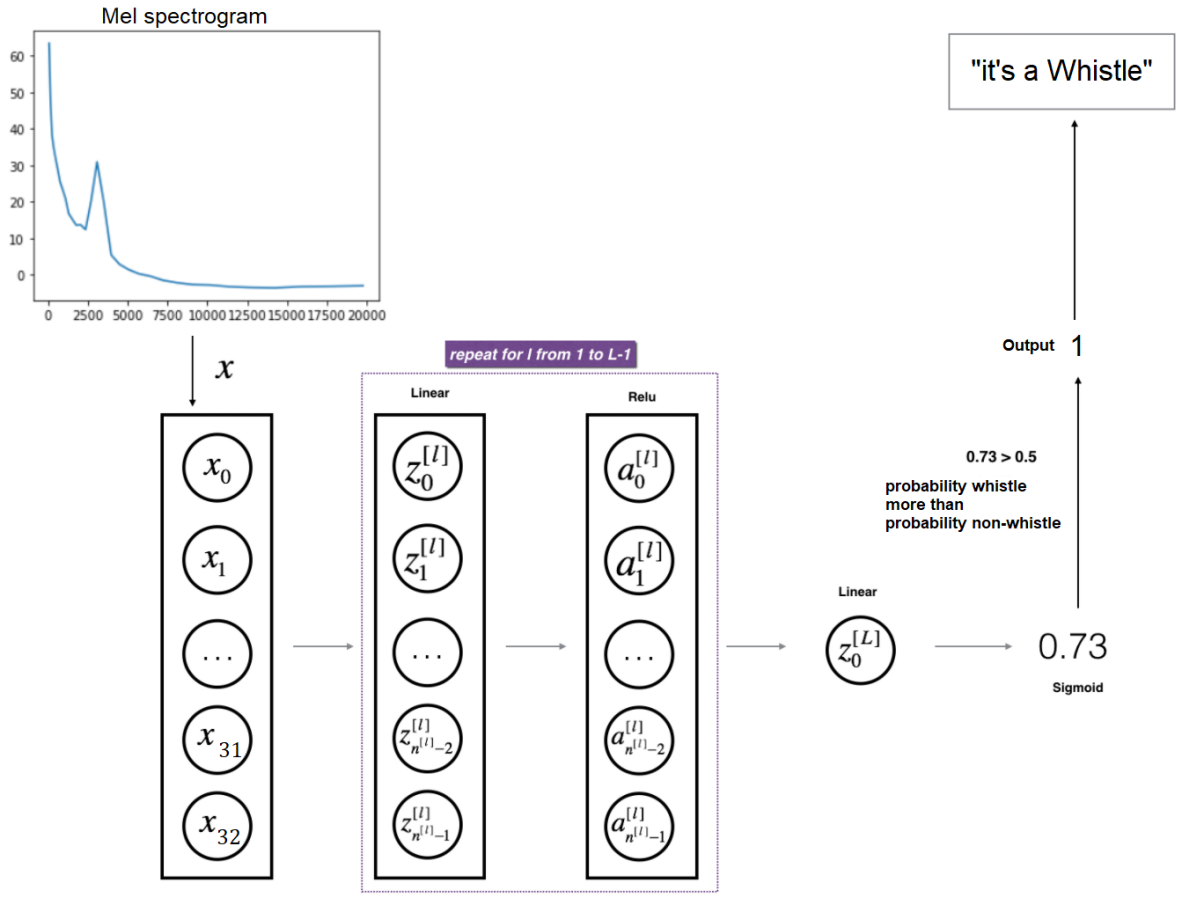


Figure 28 A simplified Whistle DNN representation. The 32 Mel frequencies are the input features of the DNN. The NN is composed by (L-1) Linear-ReLU functions plus one Linear-Sigmoid function. When the result from the sigmoid function is higher than 0.5 means that the probability that the input is Whistle is higher than the probability of a Non-Whistle case. Adapted from (13).

To obtain a prediction from the DNN, the weights and the biases (or *parameters*) are initialized with the *Xavier initialization* and the *zero initialization*, respectively (for more details see (13)). Then, the forward propagation module is performed: the input provides the initial information that propagates to the hidden units at each layer and produces an output. After that, to define the difference between the output predicted and the true label, the *Logarithmic Loss* is calculated. Finally, to reduce the loss function, the backward module is implemented, and the parameters are updated. This loop is repeated for a certain number of iterations, which is one of the hyperparameters that must be tuned in the DNN (Figure 29).

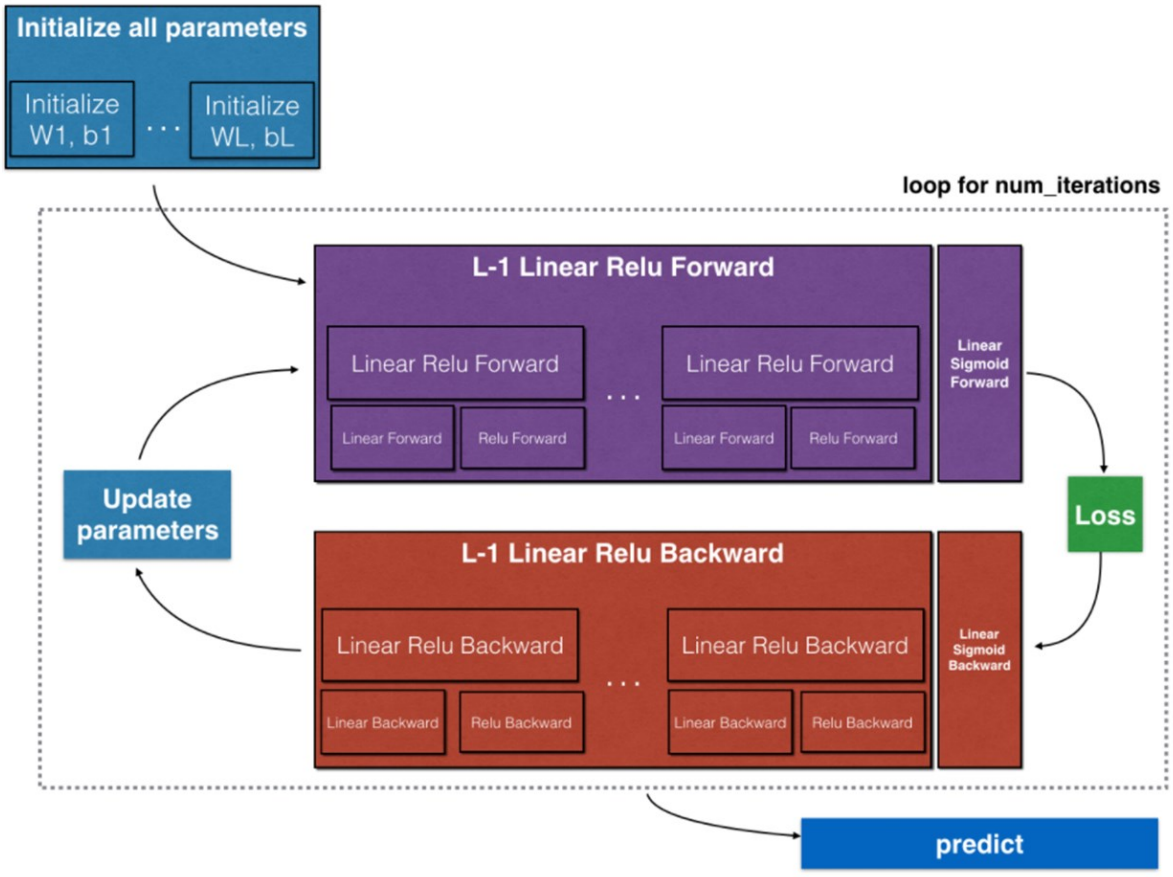


Figure 29 Schematic representation of the DNN operations. Adapted from (13).

The loss function was defined with respect to a single training example, while the *Cost function* is the mean of the loss functions and it defines how well your parameters are doing on the entire training set (for more details see (13)).

The *backward propagation* allows the information to go back from the cost backward through the network in order to compute the *gradient descent algorithm*, which permits to learn the parameters on the training set (13). For every forward function, there is a corresponding backward function (Figure 30). So, at every step of the forward module, some values have to be stored in a cache and then used to calculate the gradients.

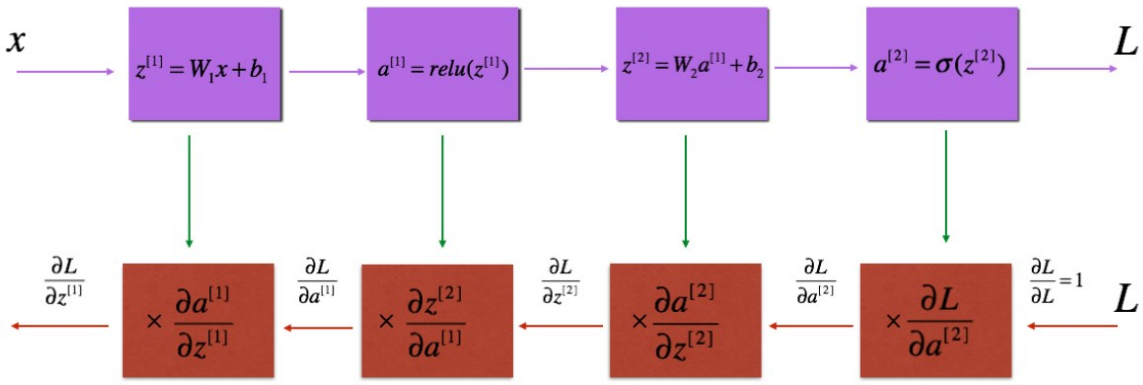


Figure 30 Relation between forward propagation (purple) and backward propagation (red). Adapted from (13)

5 Neural Networks modeling results

There are many decisions to make during the training of a NN, such as the number of hidden layers, the number of hidden units, the learning rate, the number of iterations, etc. Therefore, it is almost impossible to correctly guess the right values for all of them on the first attempt. In practice, applied ML is a highly iterative process, in which one usually starts with an idea, then runs the code and gets the accuracies, which represent a feedback about the DNN performances. Based on the outcome, a new attempt can be tried in order to find a better NN (for more details see (13)). In the following chapters, the most interesting results, obtained from this iterative procedure, are reported and discussed.

5.1 Leakage Neural Network tuning

5.1.1 Leakage Normalization

To improve NN performances, the input features normalization is widely utilized in ML because it helps the gradient descent to reduce the cost function (for more details see (13)). Therefore, the train, validation and test sets were normalized with the mean and the standard deviation calculated for the train set (for more details see Chapter 2.6.1). In Table 1, an example of the performance improvement of the same Leakage NN due to the normalization inputs is reported. A 4-layer model with the train and validation sets unnormalized achieved a train accuracy 71.0% and a validation accuracy of 73.2%, while the same 4-layer NN with the train and validation sets normalized can reach a train accuracy of 99.94% and a validation accuracy of 100%.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	NO
Number of Layers	4 - layer model	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]	[32, 20, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	3000	3000
Train set dimension	16785	16785
Validation set dimension	500	500

Train Accuracy	99.94%	71.00%
Validation Accuracy	100.00%	73.20%
Number of mislabeled cases	0	134
Percentage of mislabeled cases	0.00%	26.80%

Table 1 Comparison between Leakage NN performances with inputs normalization and without inputs normalization.

5.1.2 Number of hidden layers and hidden units

The number of hidden layers and the number of hidden units are two parameters that must be tuned to obtain a better NN. To avoid overfitting and underfitting problems, it is usually recommended to have comparable dimensions between the database and the NN capacity. In the examples shown in Table 2, the number of training cases and the number of equations are indeed comparable: 20285 cases and 22400 equations. The performances of the Leakage NN change a bit with a 3-layer model and a 4-layer model with a lower number of nodes, but they are still high.

Train/Validation Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Input Normalization	YES	YES	YES
Number of Layers	4 - layer model	3 - layer model	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]	[32, 20, 7, 1]	[32, 10, 5, 2, 1]
Learning rate	0.0075	0.0075	0.0075
Number of iterations	4500	3700	1300
Train set dimension	20285	20285	20285
Validation set dimension	500	500	500

Train Accuracy	99.97%	99.85%	99.67%
Validation Accuracy	100.00%	99.80%	99.80%
Number of mislabeled cases	0	1	1
Percentage of mislabeled cases	0.00%	0.20%	0.20%

Test Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Test set dimension	502	502	502
Test accuracy	99.80%	99.60%	99.80%
Number of mislabeled cases	1	2	1
Percentage of mislabeled cases	0.20%	0.40%	0.20%

Table 2 Comparison of Leakage NN with different number of layers and hidden units.

The shape of the 4-layer cost function with the higher number of hidden units (Figure 31) differs from the cost function of the 3-layer model (Figure 32). Both can be equally considered acceptable because they have a decreasing behavior. The cost function of the 4-layer model decreases quickly during the first iterations and then seems to stabilize with values near to zero. While, the 3-layer cost function decreases slowly during the entire running.

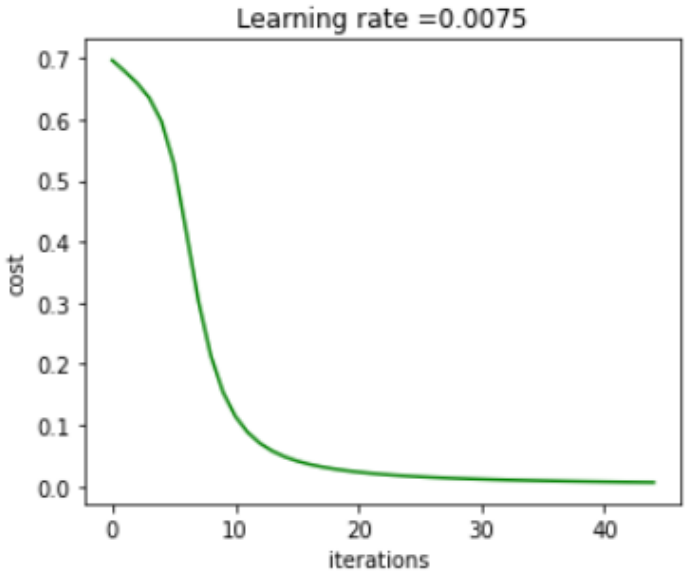


Figure 31 Cost function of the Leakage NN with 4-layer and [32,20,7,5,1] hidden units.

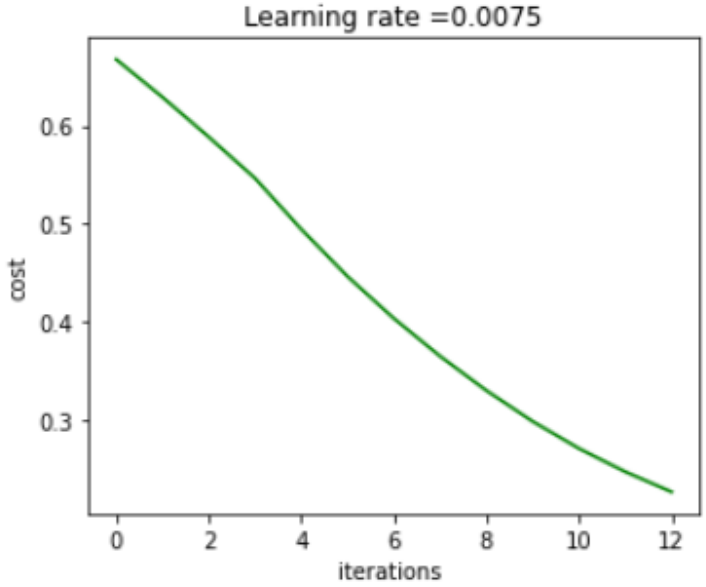


Figure 32 Cost function of the Leakage NN with 3-layer.

5.1.3 Leakage Database dimensions

Such as the number of layers and their dimensions, the training set can also change the performances of the NN due to the overfitting and underfitting considerations previously described. The train, validation and test accuracies were improved with a bigger train set and a smaller validation set. The case with the best accuracies was chosen for the following analysis, even if the other two cases have still high performances that could be considered more than acceptable (Table 3).

Train/Validation Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Input Normalization	YES	YES	YES
Number of Layers	4 - layer model	4 - layer model	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]	[32, 20, 7, 5, 1]	[32, 20, 7, 5, 1]
Learning rate	0.0075	0.0075	0.0075
Number of iterations	4500	4500	3000
Train set dimension	20285	19785	19000
Validation set dimension	500	1000	1785
Train Accuracy	99.97%	99.97%	99.91%
Validation Accuracy	100.00%	99.90%	99.89%
Number of mislabeled cases	0	0	2
Percentage of mislabeled cases	0.00%	0.00%	0.11%
Test Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Test set dimension	502	502	502
Test accuracy	99.80%	99.80%	99.80%
Number of mislabeled cases	1	1	2
Percentage of mislabeled cases	0.20%	0.20%	0.40%

Table 3 Comparison of Leakage NNs with different train and validation sets dimensions.

5.2 Performances of the Leakage Neural Network

As a result of the tuning procedure, the Leakage NN reported in Table 4 was chosen for additional analyses. Its cost function is shown in Figure 33.

Train/Validation Vehicle	Opel Vectra
Input Normalization	YES
Number of Layers	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]
Learning rate	0.0075
Number of iterations	4500
Train set dimension	20285
Validation set dimension	500

Train Accuracy	99.97%
Validation Accuracy	100.00%
Number of mislabeled cases	0
Percentage of mislabeled cases	0.00%

Test Vehicle	Opel Vectra
Test set dimension	502
Test accuracy	99.80%
Number of mislabeled cases	1
Percentage of mislabeled cases	0.20%

Table 4 The Leakage NN chosen after the tuning procedure.

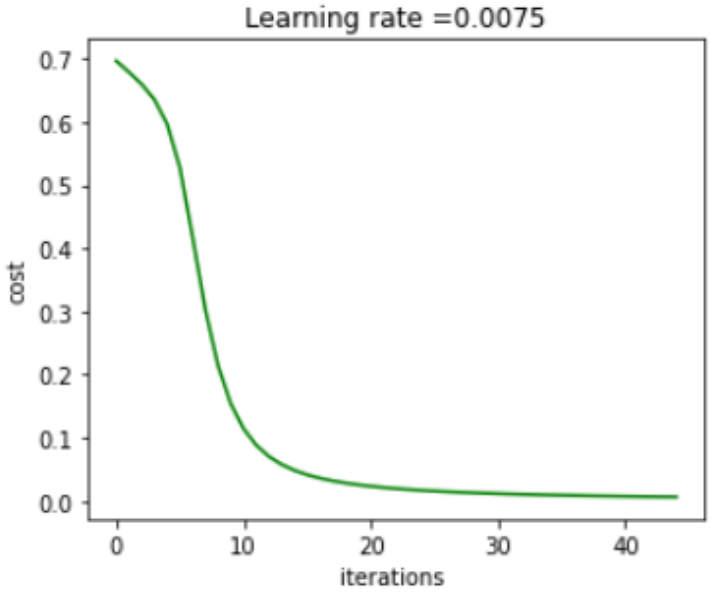


Figure 33 Cost function of the Leakage NN chosen after the tuning procedure.

When the Train, Validation and Test sets are made from the same vehicle, the accuracy on the test set is 99.8% and the Leakage NN can rightly predict the test cases (see Figure 34 and Figure 35).

```
y = 1, your L-layer model predicts a "Leakage" case.
real solution : y=1.0, there is Leakage
```

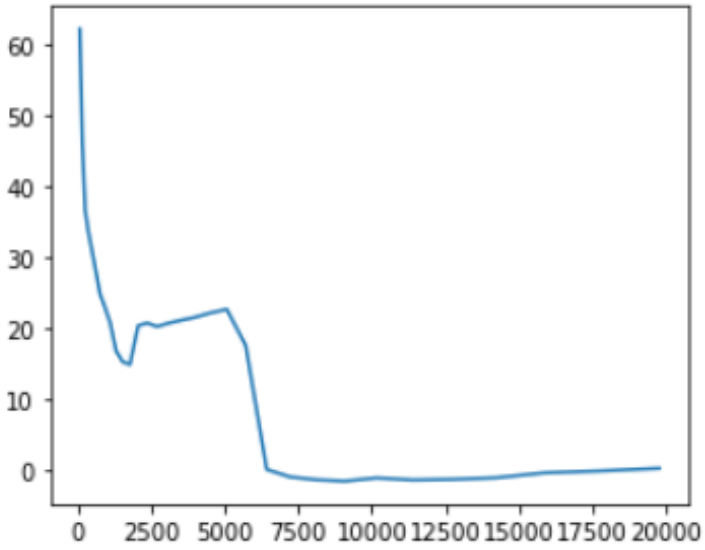


Figure 34 The Leakage NN can correctly predict the Mel spectrum with leakage.

```
y = 0, your L-layer model predicts a "Non-Leakage" case.
real solution : y=0.0, there is Non-Leakage
```

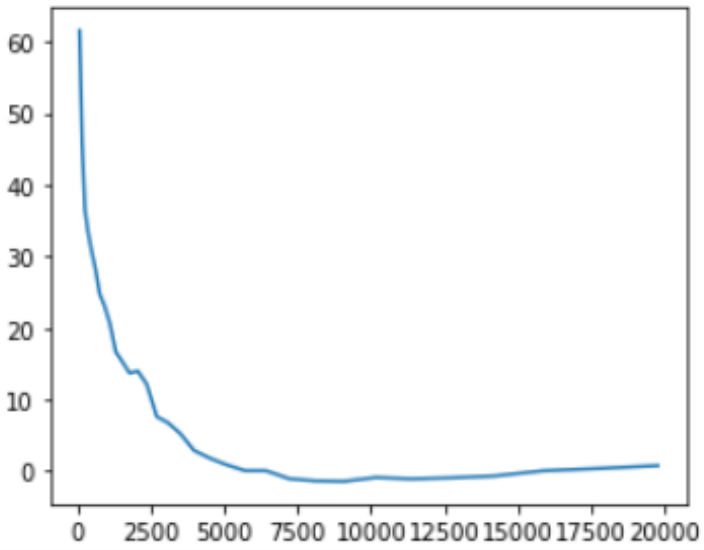


Figure 35 The Leakage NN can correctly predict the Non-Leakage Mel spectrum.

5.2.1 Train, Validation and Test sets from different vehicles

When the Train, Validation and Test sets are made from different vehicles, the Leakage NN cannot predict the presence of Non-Leakage cases (Figure 36 and Figure 37). In the example reported in Table 5, all the 250 Non-Leakage Ford Mondeo cases were mislabeled by the Leakage NN trained on Opel Vectra databases. This result was unsurprising because in supervised NNs the test and the validation sets should be made from the same distribution.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	YES
Number of Layers	4 - layer model	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]	[32, 20, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	4500	4500
Train set dimension	20285	20285
Validation set dimension	500	500

Train Accuracy	99.97%	99.97%
Validation Accuracy	100.00%	100.00%
Number of mislabeled cases	0	0
Percentage of mislabeled cases	0.00%	0.00%

Different Test and Train/Validation Vehicles		
Test Vehicle	Opel Vectra	Ford Mondeo
Test set dimension	502	502
Test accuracy	99.80%	50.19%
Number of mislabeled cases	1	250
Percentage of mislabeled cases	0.20%	49.80%

Table 5 The Leakage NN, trained on Opel Vectra train and validation sets, cannot predict the test set made from a Ford Mondeo.

```
y = 1, your L-layer model predicts a "Leakage" case.
real solution : y=1.0, there is Leakage
```

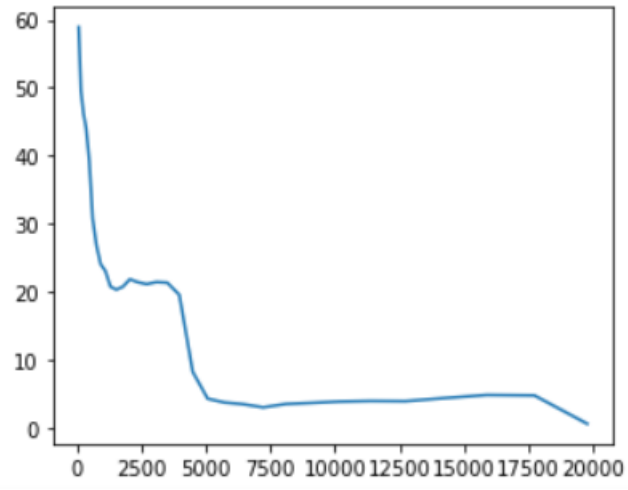


Figure 36 The Leakage NN, trained on Opel Vectra train and validation sets, can correctly predict the Mel spectrum with leakage of a Ford Mondeo.

```
y = 1, your L-layer model predicts a "Leakage" case.
real solution : y=0.0, there is Non-Leakage
```

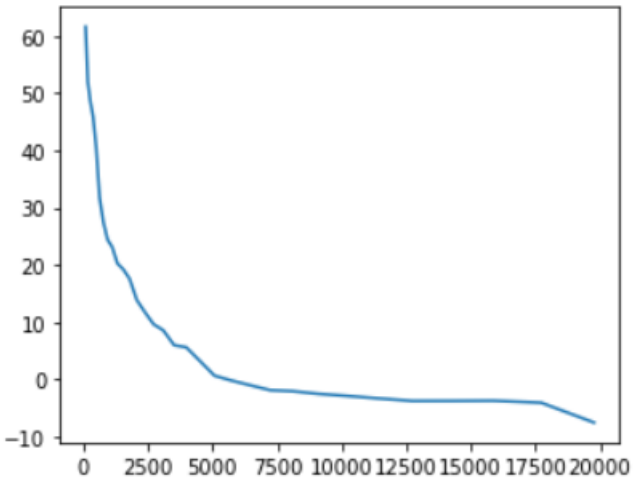


Figure 37 The Leakage NN, trained on an Opel Vectra database, cannot correctly predict the Mel spectrum without leakage of a Ford Mondeo.

5.2.2 Test case with Whistle cases labelled as Non-Leakage cases

When in the test set there are Whistle cases labeled as Non-Leakage cases (Table 6), the Leakage NN cannot predict any of the Non-Leakage cases, not even the Non-Leakage and Non-Whistle cases (Figure 38 and Figure 39). These results were expected because the validation and the test sets were made from different distribution.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	YES
Number of Layers	4 - layer model	4 - layer model
Dimensions of the layers	[32, 20, 7, 5, 1]	[32, 20, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	4500	4500
Train set dimension	20285	20285
Validation set dimension	500	500

Train Accuracy	99.97%	99.97%
Validation Accuracy	100.00%	100.00%
Number of mislabeled cases	0	0
Percentage of mislabeled cases	0.00%	0.00%

Test Database with Whistle cases labelled as Non-Leakage cases		
Test Vehicle	Opel Vectra	Opel Vectra
Test set dimension	502	502
Test accuracy	99.80%	0.00%
Number of mislabeled cases	1	502
Percentage of mislabeled cases	0.20%	100.00%

Table 6 Test database with Whistle cases labelled as Non-Leakage cases cannot be correctly predicted by the Leakage NN.

```
y = 1, your L-layer model predicts a "Leakage" case.
real solution : y=0.0, there is Non-Leakage
```

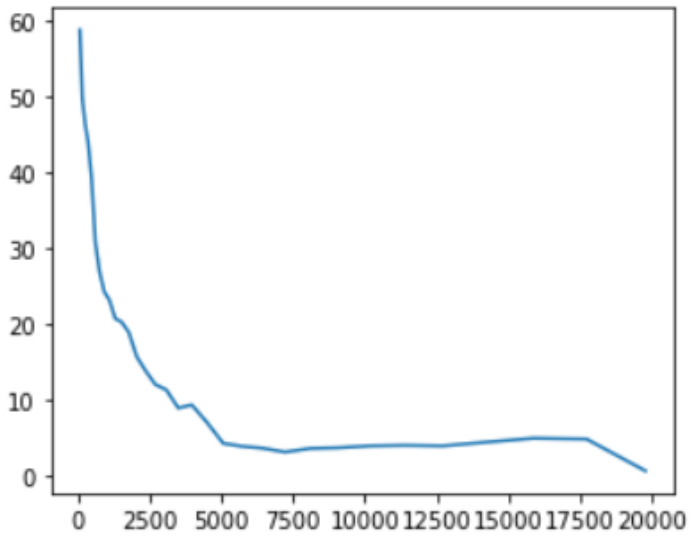


Figure 38 The Leakage NN, trained on Leakage and Non-Leakage and Non-Whistle cases, cannot predict the Non-Leakage_Non-Whistle cases made from a different distribution.

```
y = 1, your L-layer model predicts a "Leakage" case.
real solution : y=0.0, there is Non-Leakage
```

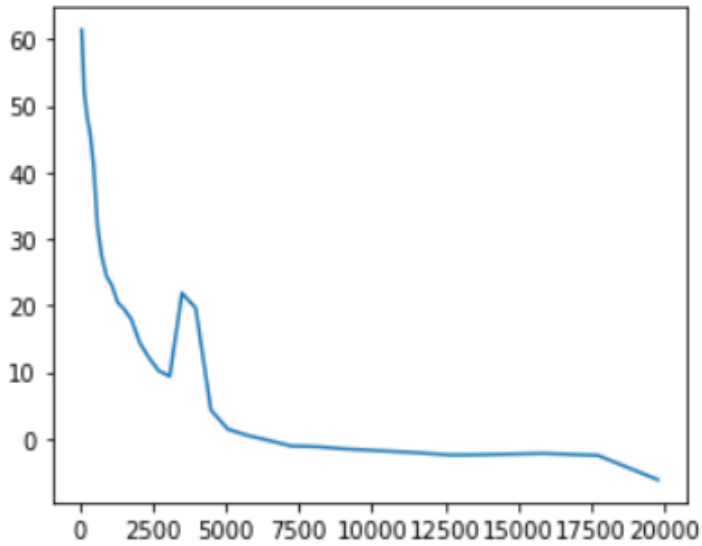


Figure 39 The Leakage NN cannot predict the Whistle cases labelled as Non-Leakage cases.

5.3 Whistle Neural Network tuning

All the previous considerations made about the Leakage NN tuning, can be also formulated for the Whistle NN.

5.3.1 Whistle Normalization

In Table 7, an example of performance improvements of the same Whistle NN due to the normalization inputs is reported. A 6-layer model with the train and validation sets unnormalized achieved a train accuracy 59.85% and a validation accuracy of 61.40%, while the same 6-layer NN with the train and validation sets normalized can reach a train accuracy of 88.90% and a validation accuracy of 90.10%. These values are lower than the ones reached by the Leakage NN, but still acceptable.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	NO
Number of Layers	6 - layer model	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	3800	3800
Train set dimension	11504	11504
Validation set dimension	1000	1000

Train Accuracy	88.90%	59.85%
Validation Accuracy	90.10%	61.40%
Number of mislabeled cases	99	386
Percentage of mislabeled cases	9.90%	38.60%

Table 7 Comparison between Leakage DNN performances with inputs normalization and without inputs normalization.

The shape of the Whistle NN cost function changes before and after the inputs' normalization. Both can be equally considered acceptable because they have a decreasing behavior. The cost function of Whistle NN without normalization decreases quickly during the first iterations and then seems to stabilize with values between 0.6725 and 0.6750 (Figure 40). While, the cost function after the normalization remains constant at the beginning, then it decreases quickly at the end of the training and reaches lower values (Figure 41).

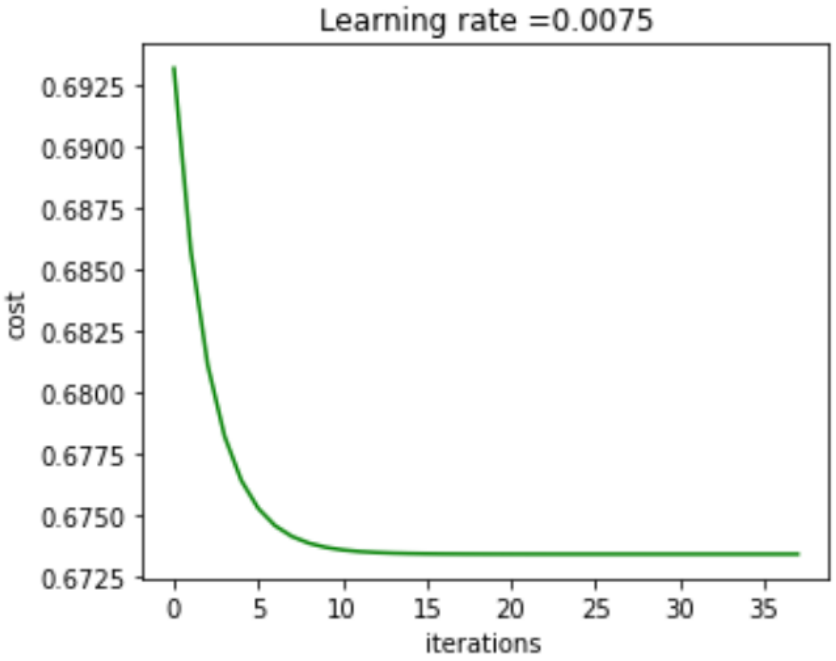


Figure 40 Cost function of the Whistle NN before the input normalization.

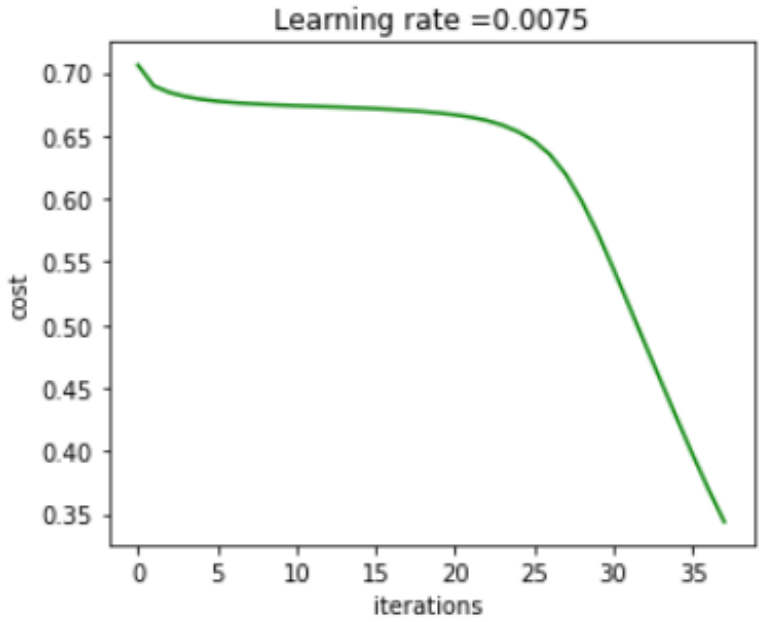


Figure 41 Cost function of the Whistle NN after the input normalization.

5.3.2 Number of hidden layers and hidden units

As already mentioned before, to avoid overfitting and underfitting problems, it is usually recommended to have comparable dimensions between the database and the NN capacity. However, in the examples shown in Table 8, the NN with the best performances have the number of equations that are three orders higher than the number of training cases. In ML is not unusual to find NNs with high number of hidden layers and units that perform better than the others because the NN tends to use only the nodes that are necessary for its training.

Train/Validation Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Input Normalization	YES	YES	YES
Number of Layers	6 - layer model	4 - layer model	4 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]	[32, 20,10, 5, 1]	[32, 20,7, 5, 1]
Learning rate	0.0075	0.0075	0.0075
Number of iterations	3800	1800	2800
Train set dimension	19932	19932	19932
Validation set dimension	1000	1000	1000
Train Accuracy	88.79%	86.10%	87.35%
Validation Accuracy	90.20%	88.80%	89.50%
Number of mislabeled cases	98	112	105
Percentage of mislabeled cases	9.80%	11.20%	10.50%
Test Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Test set dimension	1028	1028	1028
Test accuracy	92.80%	89.10%	92.90%
Number of mislabeled cases	74	112	73
Percentage of mislabeled cases	7.20%	10.89%	7.10%

Table 8 Comparison of Leakage DNN with different number of layers and hidden units.

The shape of the 4-layer cost function (Figure 42) differs from the cost function of the 6-layer model (Figure 43). The cost function of the 4-layer model decreases slowly during the entire running and so it can still be considered acceptable.

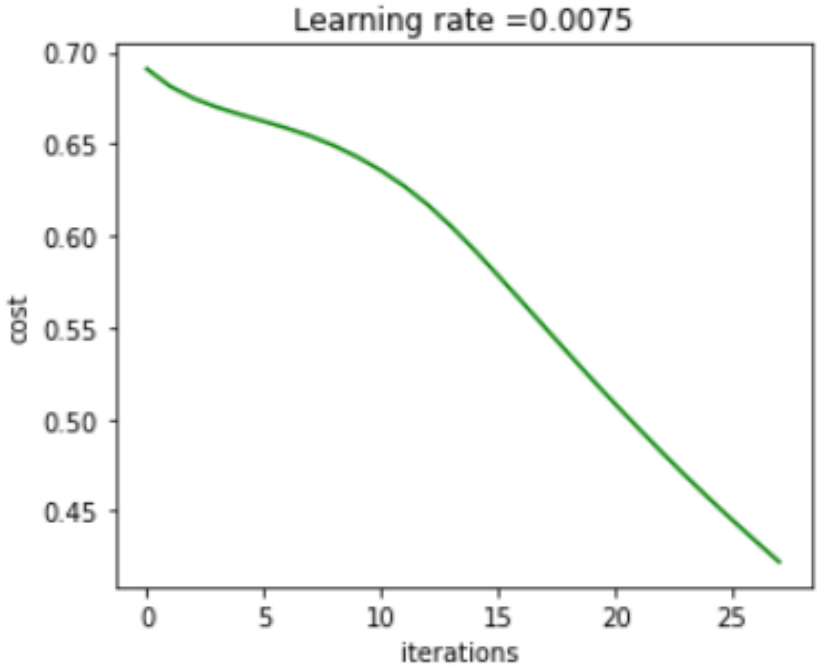


Figure 42 Cost function of the Whistle NN with 4 layers.

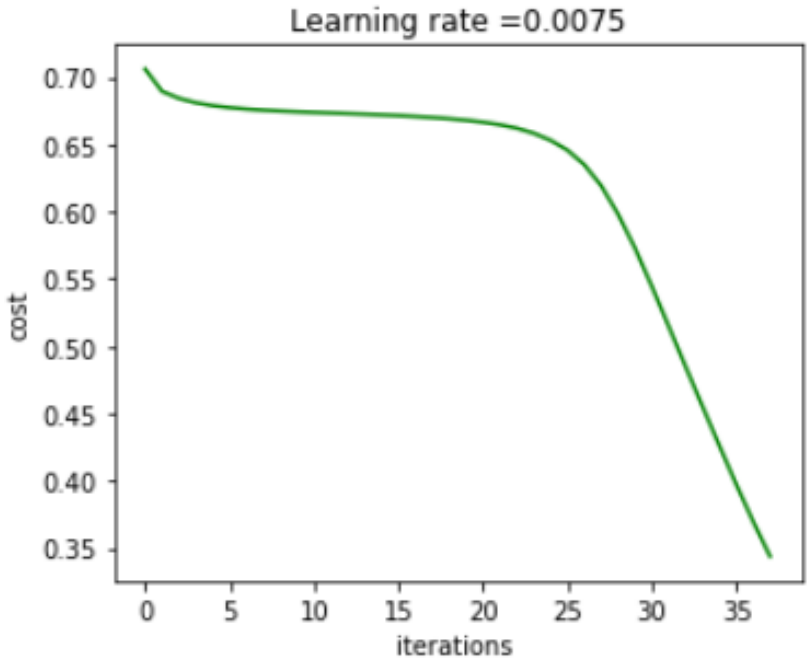


Figure 43 Cost function of the Whistle NN with 6 layers.

5.3.3 Whistle Database dimensions

Such as the number of layers and their dimensions, the training set dimension can change the performances of the NN due to the overfitting and underfitting considerations, previously described. The case with the best accuracies was chosen for the following considerations (Table 9).

Train/Validation Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Input Normalization	YES	YES	YES
Number of Layers	6 - layer model	6 - layer model	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]	[32, 30, 25, 15, 7, 5, 1]	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075	0.0075	0.0075
Number of iterations	3800	3800	3800
Train set dimension	19932	19432	20432
Validation set dimension	1000	1500	500
Train Accuracy	88.79%	88.73%	88.57%
Validation Accuracy	90.20%	88.47%	89.60%
Number of mislabeled cases	98	173	52
Percentage of mislabeled cases	9.80%	11.53%	10.40%
Test Vehicle	Opel Vectra	Opel Vectra	Opel Vectra
Test set dimension	1028	1028	1028
Test accuracy	92.80%	92.50%	92.41%
Number of mislabeled cases	74	77	78
Percentage of mislabeled cases	7.20%	7.49%	7.59%

Table 9 Comparison of Whistle NNs with different train and validation sets dimensions.

5.4 Performances of the Whistle Neural Network

As a result of the tuning procedure, the Whistle NN reported in Table 10 was chosen for additional analyses. The accuracies reached by the Whistle NN are lower than the one reached by the Leakage NN. However, they are still good respect to the human operator which could realize around $60\% - 70\%$ of accuracy. The cost function is shown in Figure 44.

Train/Validation Vehicle	Opel Vectra
Input Normalization	YES
Number of Layers	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075
Number of iterations	3800
Train set dimension	19932
Validation set dimension	1000

Train Accuracy	88.79%
Validation Accuracy	90.20%
Number of mislabeled cases	98
Percentage of mislabeled cases	9.80%

Test Vehicle	Opel Vectra
Test set dimension	1028
Test accuracy	92.80%
Number of mislabeled cases	74
Percentage of mislabeled cases	7.20%

Table 10 The Whistle NN chosen after the tuning procedure.

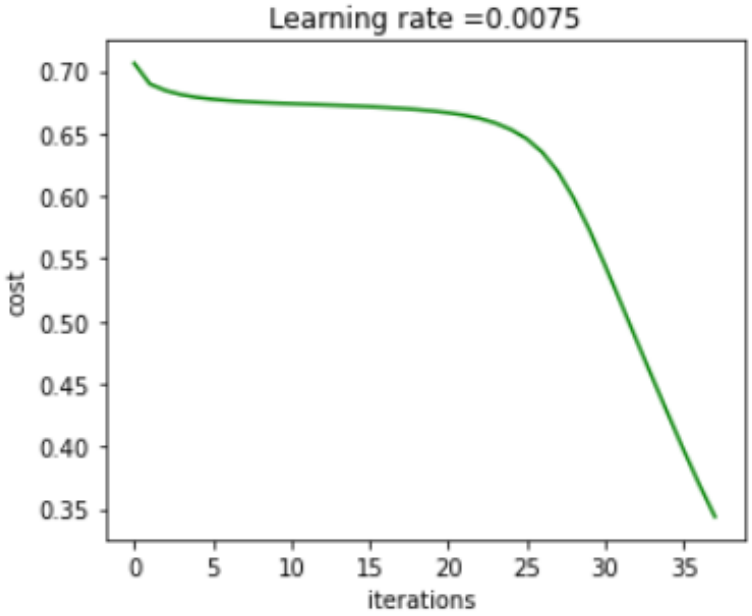


Figure 44 Cost function of the Whistle NN chosen after the tuning procedure.

When the Train, Validation and Test sets are made from the same vehicle, the accuracy on the test set is 92.8% and the Whistle NN can rightly predict the test cases (see Figure 45 and Figure 46).

```
y = 1, your L-layer model predicts a "Whistle" case.
real solution : y=1.0, it is a "Whistle" case.
```

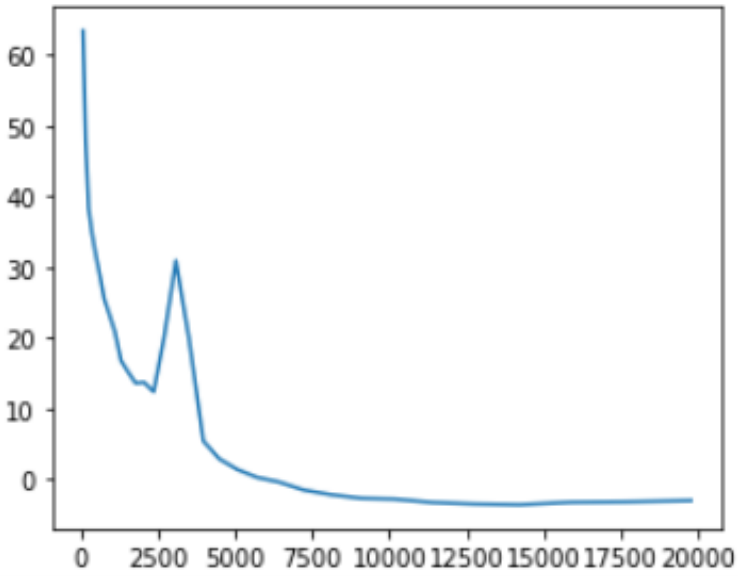


Figure 45 The Whistle NN can correctly predict the Mel spectrum with whistle.

```
y = 0, your L-layer model predicts a "Non-Whistle" case.
real solution : y=0.0, it is a "Non-Whistle" case.
```

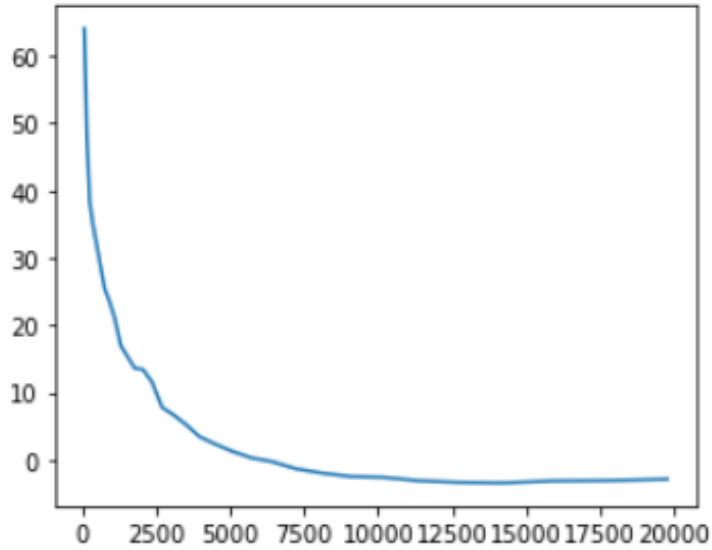


Figure 46 The Whistle NN can correctly predict the Non-Whistle Mel spectrum.

5.5.1 False positive e false negative cases

After an analysis of the 98 mislabeled cases in the Validation set, 33 cases resulted as False Positive and 0 case as False Negative. There are also 35 mislabeled cases that have a small peak, which were arduous to predict by the Whistle NN (Table 11).

Train/Validation Vehicle	Opel Vectra
Input Normalization	YES
Number of Layers	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075
Number of iterations	3800
Train set dimension	19932
Validation set dimension	1000

Train Accuracy	88.79%
Validation Accuracy	90.20%
Number of mislable cases	98
Percentage of mislable cases	9.80%

Number of False Positive	33
Percentage of False Positive	3.30%
Number of Positive with small peak	35
Percentage of Positive with small peak	3.50%
Number of False Negative	0

Table 11 Number of False Positive and False Negative cases in the Validation set.

Two examples of “False Positive” and “Positive with small peak” are shown in Figure 47 and Figure 48, respectively. As can be seen from the following figures, these cases are the results of small peaks at low frequencies that were covered up by background noises. They could be one of the reasons of the Whistle NN lower performances. However, the ranges values selected during the Whistle generation were the result of experimental measures and physical consideration (for more details see Chapter 3.4). Thus, the most effective way, to avoid these cases and don’t change the values ranges, could be to develop an unsupervised NN, such as an autoencoder, that could discern the heathy from the unhealthy cases. In this way, these cases would be considered healthy and only the unhealthy cases would be analyzed by the DNN.

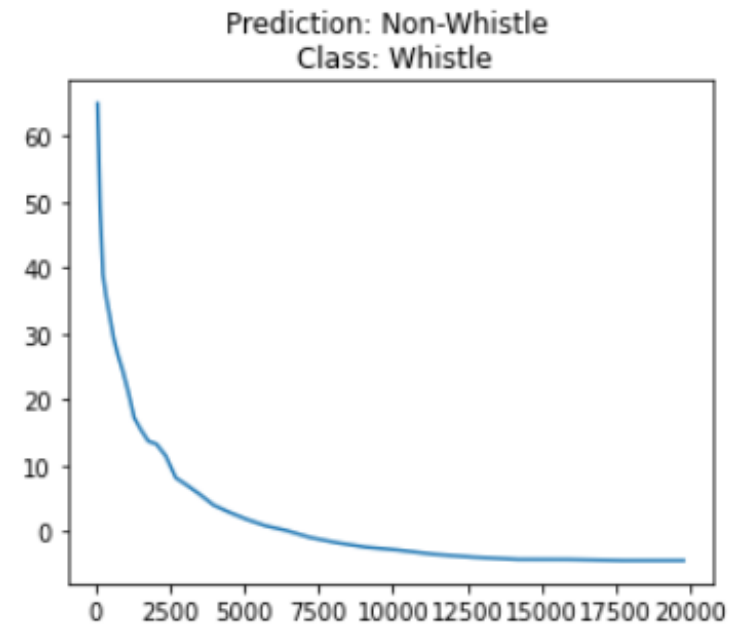


Figure 47 Example of False Positive case.

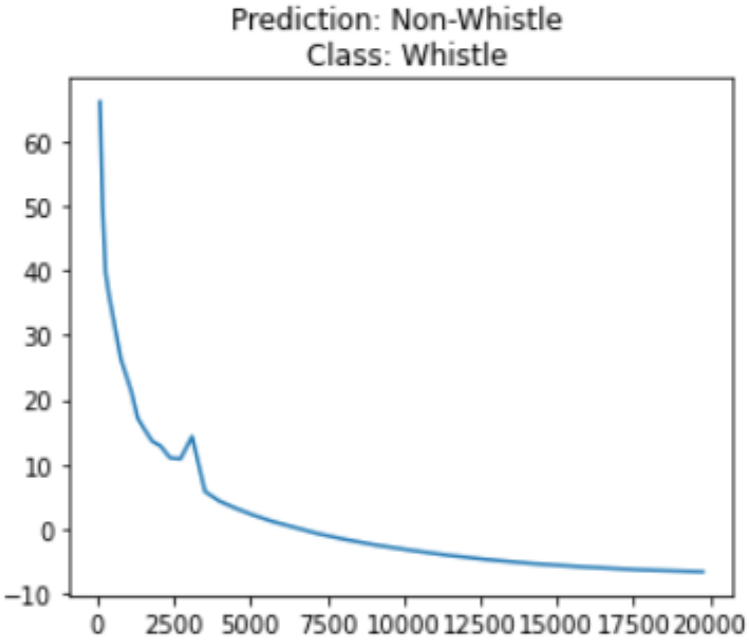


Figure 48 Example of Positive case with a small Whistle peak.

5.4.1 Train, Validation and Test sets from different vehicles

As previously mentioned, in supervised NNs the test and the validation sets should be made from the same distribution. Thus, it was unsurprising that when the Train, Validation and Test sets were made from different vehicles, the Whistle NN could not predict the presence of Non-Whistle cases (Figure 49 and Figure 50). In the example reported in Table 12, all the 500 Non-Whistle Ford Mondeo cases were mislabeled by the Whistle NN trained on Opel Vectra databases.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	YES
Number of Layers	6 - layer model	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	3800	3800
Train set dimension	19932	19932
Validation set dimension	1000	1000

Train Accuracy	88.79%	88.79%
Validation Accuracy	90.20%	90.20%
Number of mislabeled cases	98	98
Percentage of mislabeled cases	9.80%	9.80%

Different Test and Train/Validation Vehicles		
Test Vehicle	Opel Vectra	Ford Mondeo
Test set dimension	1028	1028
Test accuracy	92.80%	51.36%
Number of mislabeled cases	74	500
Percentage of mislabeled cases	7.20%	48.64%

Table 12 The Whistle NN, trained on Opel Vectra train and validation sets, cannot predict the test set made from a Ford Mondeo.

```
y = 1, your L-layer model predicts a "Whistle" case.
real solution : y=0.0, it is a "Non-Whistle" case.
```

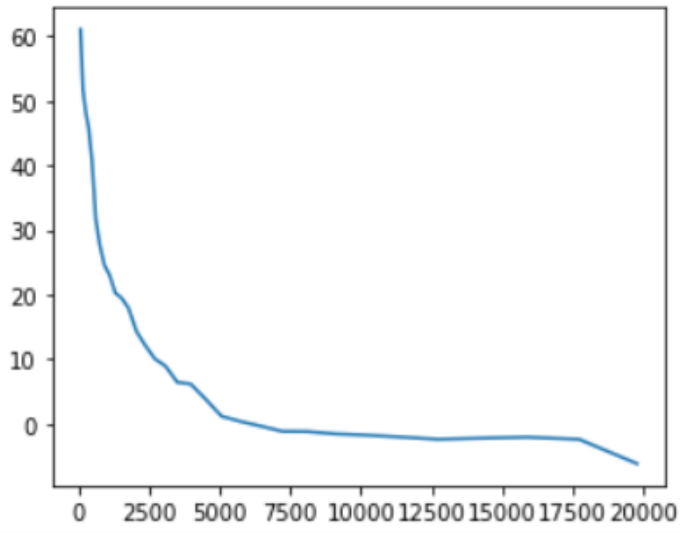


Figure 49 The Whistle NN, trained on an Opel Vectra database, cannot correctly predict the Mel spectrum without whistle of a Ford Mondeo.

```
y = 1, your L-layer model predicts a "Whistle" case.
real solution : y=1.0, it is a "Whistle" case.
```

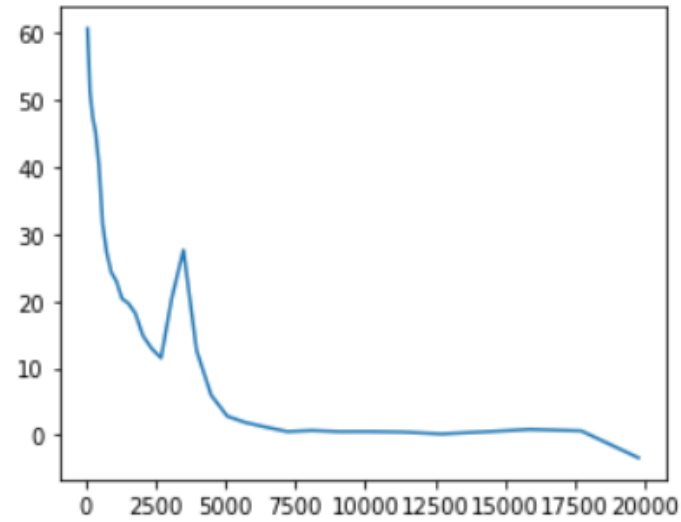


Figure 50 The Whistle NN, trained on a Opel Vectra, can correctly predict the Mel spectrum with whistle of a Ford Mondeo.

5.4.2 Test case with Leakage cases labelled as Non-Whistle cases

When in the test set there are Leakage cases labeled as Non-Whistle cases (Table 13), the Whistle NN cannot predict any of the Non-Whistle cases, not even the Non-Leakage and Non-Whistle cases (Figure 51 and Figure 52). These results were expected because, as mentioned previously, the validation and the test sets were made from different distribution.

Train/Validation Vehicle	Opel Vectra	Opel Vectra
Input Normalization	YES	YES
Number of Layers	6 - layer model	6 - layer model
Dimensions of the layers	[32, 30, 25, 15, 7, 5, 1]	[32, 30, 25, 15, 7, 5, 1]
Learning rate	0.0075	0.0075
Number of iterations	3800	3800
Train set dimension	19932	19932
Validation set dimension	1000	1000

Train Accuracy	88.79%	88.79%
Validation Accuracy	90.20%	90.20%
Number of mislabeled cases	98	98
Percentage of mislabeled cases	9.80%	9.80%

Test Database with Leakage cases labeled as Non-Whistle cases		
Test Vehicle	Opel Vectra	Opel Vectra
Test set dimension	1028	1028
Test accuracy	92.80%	0.00%
Number of mislabeled cases	74	1028
Percentage of mislabeled cases	7.20%	100.00%

Table 13 Test database with Leakage cases labelled as Non-Whistle cases cannot be correctly predicted by the Whistle NN.

```
y = 1, your L-layer model predicts a "Whistle" case.
real solution : y=0.0, it is a "Non-Whistle" case.
```

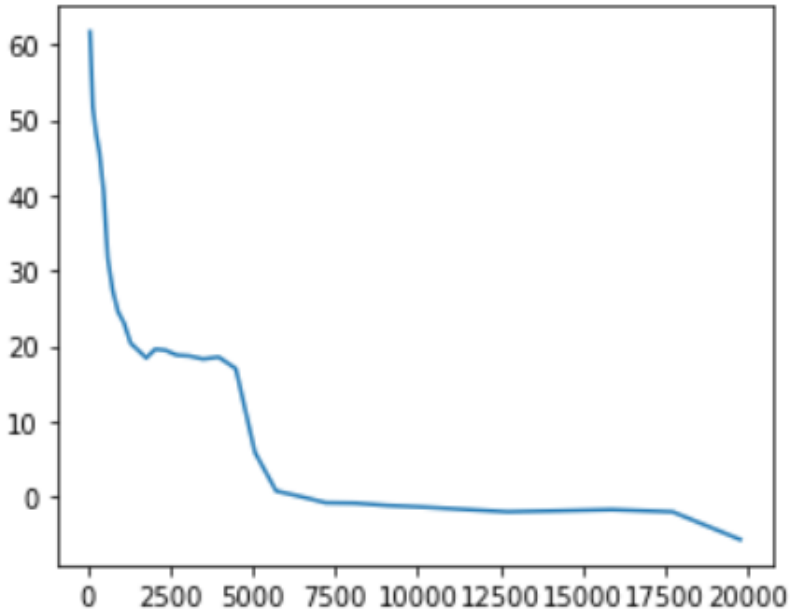


Figure 51 The Whistle NN cannot predict the Leakage cases labelled as Non-Whistle cases.

```
y = 1, your L-layer model predicts a "Whistle" case.
real solution : y=0.0, it is a "Non-Whistle" case.
```

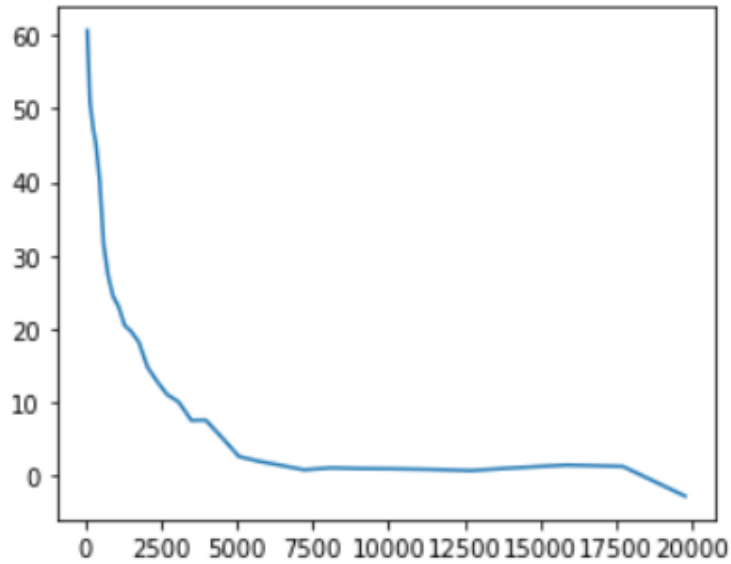


Figure 52 The Whistle NN, trained on Whistle and Non-Leakage_Non-Whistle cases, cannot predict the Non-Leakage_Non-Whistle cases made from a different distribution.

6 Conclusions and Recommendations

Two different deep connected Neural Networks were implemented for the detection of leakages and whistling noises in a full vehicle End-of-Line. The Leakage NN achieves the highest performances: 99.97% of Train accuracy, 100% of Validation accuracy and 99.80% of Test accuracy. The Whistle NN reaches 88.79% of Train accuracy, 90.20% of Validation accuracy and 92.80% of Test accuracy. There are 98 mislabelled cases in the Validation set of the Whistle NN, whose 33 cases are False Positives and 35 cases have a small peak that could be hard for the NN to detect. The Whistle NN has lower performances but still good respect to the human operator which could realize around $60\% - 70\%$ of accuracy.

One of the DNNs' weaknesses is their necessity to be trained with a database that has the same distribution of the Test set. Indeed, when the Test set is composed by different types of vehicle or there are Whistle cases labelled as Non-Leakage cases for the Leakage NN or Leakage cases labelled as Non-Whistle cases for the Whistle NN, both NNs cannot rightly predict the presence of whistles and leakages. It's possible to generalize the train and validation databases with background noises from different vehicles to improve the NNs flexibility. To solve the mislabelled cases problem, Whistle cases classified as Non-Leakage and Leakage cases labelled as Non-Whistle could be added to the Leakage database and Whistle database, respectively. Another possibility could be to try a DNN with three outputs that could detect Leakage, Whistle and Non-Leakage_Non-Whistle cases, or different structures such as CNN. An alternative option is to add an unsupervised NN, such as autoencoders, that could discern the healthy from the unhealthy cases and then use a DNN, like the ones implemented in this thesis, that could distinguish the Leakages from the Whistles. These suggestions are just few solutions that could be developed in the future. The methodology described in this thesis could be considered as an encouragement for further researches in Machine Learning, which holds great potential for identifying and quantifying NVH issues by automated procedures.

Bibliography

1. *Vehicle Interior Noise - Combination of Sound, Vibration and Interactivity.* **Genuit, Klaus.** 2009, Vols. Sound & vibration. 43. 8-13.

2. **Helfer, M.** *GENERAL ASPECTS OF VEHICLE AEROACOUSTICS.* 2005.

3. **Michael J. Bianco, Peter Gerstoft, James Traer, Emma Ozanich, Marie A. Roch,.** *Machine learning in acoustics: Theory and applications.* 2019.

4. **West B., Kendrick P.** *Automotive Aeroacoustic Sound Quality.* 2019.

5. *Automotive aeroacoustics : an overview.* **Oettle, Nicholas and Sims-Williams, David.** 2017, Vols. Proceedings of the Institution of Mechanical Engineers. Part D : Journal of automobile engineering., 231 (9). pp. 1177-1189.

6. *Source Decomposition for Vehicle Sound Simulation.* **M. Allman-Ward, M.P. Balaam and R. Williams,.** s.l. : CETIM Conference, 2001.

7. **Alun Crewe, David Bogema, Roger Williams, Murray Balaam, Mark Allman-Ward.** *Sound Decomposition - A Key to Improved Sound Simulation.* 2003.

8. **What's an Order?** ***Siemens Community.*** **[Online] 7 October 2020. https://community.sw.siemens.com/s/article/what-s-an-order.**

9. Blough, Jason R., David L. Brown, and Håvard Vold. ***The time variant discrete Fourier transform as an order tracking method.*** **s.l. : SAE transactions: 3037-3045., 1997.**

10. Leuridan, Vold Håvard and Jan. ***High resolution order tracking at extreme slew rates, using Kalman tracking filters.*** **s.l. : No. 931288. SAE Technical Paper, 1993.**

11. ***Synthesis techniques for wind and tire-road noise.*** **Sarrizin, M., Colangeli, C., Janssens, K., & van der Auweraer, H. s.l. : In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 247, No. 6, pp. 2303-2312). Institute of Noise Control Engineering, 2013, September. Vols. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 247, No. 6, pp. 2303-2312). Institute of Noise Control Engineering.**

12. Roberts, Leland. Understanding the Mel Spectrogram. ***Analytics vidhya.*** **[Online] 6 March 2020. https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53.**

13. Andrew Ng, Kian Katanforoosh, Younes Bensouda Mourri. Deep Learning Specialization. ***Coursera.org.*** **[Online]**

14. Roberts, Leland. Understanding the Mel Spectrogram. ***Analytics Vidhaya.*** **[Online] 6 March 2020. https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53.**