Triage for out-of-hospital emergency incidents represents a tough challenge, primarily due to time constraints—requiring rapid priority assessment within a narrow time frame—and uncertainty—making decisions with limited available information. Furthermore, errors in this process can have severe consequences for patients, potentially leading to death. Therefore, any novel protocol, tool, or strategy that has been demonstrated to enhance these processes can offer substantial value in terms of patient care and overall management of out-of-hospital emergency medical incidents.

The fundamental hypothesis upon which this thesis is based is that Machine Learning, specifically Deep Learning, can significantly improve these processes by providing estimations of the severity of out-of-hospital emergency medical incidents, taking into account the information available to the dispatcher at the moment of incident prioritization during the emergency call. By analyzing millions of data derived from emergency calls from the Valencian Region (Spain) spanning from 2009 to 2019, we posited that Machine Learning models could extract patterns that may confer predictive capability to this task.

Hence, this thesis delves into designing and developing various Machine Learning models, specifically Deep Multitask Learning models that leverage multimodal out-of-hospital emergency medical data. Our primary objective was to predict three labels indicative of incident severity, thereby influencing its prioritization. These labels encompassed whether the incident posed a life-threatening situation, the admissible response delay (ranging from non-delayable to minutes, hours, or days), and whether it fell under the jurisdiction of the emergency system or primary care. Using data available from 2009 to 2012, the results obtained were promising. We observed substantial improvements in macro F1-scores, with gains of 12.5% for life-threatening classification, 17.5% for response delay, and 5.1% for jurisdiction classification, compared to the in-house triage protocol of the Valencian Region.

However, it is essential to note that systems, dispatch protocols, and operational practices naturally evolve over time. Models that exhibited excellent performance with the initial dataset from 2009 to 2012 did not demonstrate the same efficacy when evaluated on data spanning from 2014 to 2019 (data from 2013 were not available). This later dataset had undergone significant modifications compared to the earlier one. These modifications led to dataset shifts, resulting in variations in probability distributions, which we have meticulously characterized and investigated in this thesis, focusing on their impact on model performance.

Continuing our research, we aimed to provide sustainable model performance over time or, at the very least, to mitigate the adverse effects of the inevitable distribution variations as effectively as possible. To address this challenge, we placed our focus on Deep Continual Learning. By incorporating the Continual Learning paradigm into our designs and developments, we could substantially mitigate the adverse performance effects and better understand how to manage model deployment over time in an emergency medical dispatch center. The results of our research indicate that when considering Deep Continual Learning, while it may not entirely eliminate performance fluctuations over time, it effectively maintains them within a manageable range. In particular, with respect to the F1-score, when distributional variations fall within the light to moderate range, the performance remains stable, not varying by more than 2.5%, as observed in our out-of-hospital medical incident data. Therefore, under these conditions, our models' performance is operationally acceptable.

Furthermore, our thesis demonstrates the feasibility of building auxiliary tools that enable dispatchers to interact with these complex deep models. Consequently, without disrupting professionals' workflow, it becomes possible to provide feedback through probability predictions for each severity label class and take appropriate actions based on these predictions.

Finally, the outcomes of this thesis hold direct implications for the management of out-of-hospital emergency medical incidents in the Valencian Region. The final model resulting from our research is slated for integration into the emergency medical dispatch centers of the Valencian Region. This model will utilize data provided by dispatchers to automatically compute severity predictions, which will then be compared with those generated by the in-house triage protocol. Any disparities between these predictions will trigger the referral of the incident to a physician coordinator, who will oversee its handling. Therefore, it is evident that our thesis, in addition to making significant contributions to the field of Biomedical Machine Learning Research, also carries substantial implications for enhancing the management of out-of-hospital emergencies in the context of the Valencian Region.