

UNIVERSITÀ DEGLI STUDI DI  
MILANO-BICOCCA

ADVANCED MACHINE LEARNING  
FINAL PROJECT

---

Title

---

*Authors:*

John Smith - XSFT901121- john.smith@email.com

John Doe - MTS4128991- john.doe@email.com

January 4, 2022



## Abstract

Una immagine vale più di mille parole. Sapevi che una immagine può salvare più di mille vite? Milioni di animali randagi soffrono per le strade o vengono soppressi nei rifugi ogni giorno in tutto il mondo. È chiaro aspettarsi che gli animali con foto più "attraenti" generino più interesse e vengano adottati più velocemente. Ma cosa rende "attraente" una immagine? Con l'aiuto del Machine Learning cercheremo di determinare l'attrattiva di una foto di un animale al fine di dargli una maggiore possibilità di adozione.

(TO DELETE) The ABSTRACT is not a part of the body of the report itself. Rather, the abstract is a brief summary of the report contents that is often separately circulated so potential readers can decide whether to read the report. The abstract should very concisely summarize the whole report: why it was written, what was discovered or developed, and what is claimed to be the significance of the effort. The abstract does not include figures or tables, and only the most significant numerical values or results should be given.

## 1 Introduction

Il task fa riferimento ad una competizione aperta su Kaggle. La competizione è finanziata da PetFinder.my, principale piattaforma per il benessere degli animali della Malesia. Attualmente, PetFinder.my utilizza un misuratore di carineria per classificare le foto di animali domestici. Questo misuratore analizza la composizione della foto e altri fattori rispetto alle valutazioni di migliaia di profili di animali. Sebbene questo strumento sia utile, è ancora in una fase sperimentale e l'algoritmo potrebbe essere migliorato. L'obiettivo è quello di stimare la Pawpularity (termine coniato dalla piattaforma per indicare la popolarità di una foto) di un animale in base alla foto del suo profilo. Insieme alle foto di migliaia di animali vengono anche forniti dei metadati etichettati a mano per ogni foto. Vengono forniti quindi due dataset, uno con le foto degli animali e l'altro con i corrispettivi metadati. La soluzione da noi proposta utilizza un modello multi-input. Da un lato una rete convoluzione che sfrutta un approccio di fine-tuning basato su transfer learning, dall'altro una semplice rete neurale. Gli output delle due reti vengono concatenati per ottenere l'output finale, ovvero il Pawpularity score.

The introduction should provide a clear statement of the problem posed by the project, and why the problem is of interest. It should reflect the

scenario, if available. If needed, the introduction also needs to present background information so that the reader can understand the significance of the problem. A brief summary of the hypotheses and the approach your group used to solve the problem should be given, possibly also including a concise introduction to theory or concepts used later to analyze and to discuss the results.

## 2 Datasets

Vengono forniti un dataset di immagini e un dataset tabulare contenente i metadati delle immagini. In particolare: I dati di training comprendono:

- Una cartella **train** contenente le immagini di training nella forma {id}.jpg, dove id è un identificativo univoco del profilo dell'animale;
- Un file **train.csv** contenente i metadati per ogni immagine insieme al target, che è il Pawpularity score.

I dati di testing, dato che si tratta di una competizione aperta, non sono disponibili. Tuttavia, vengono fornite la cartella **test** e il file **test.csv** di otto immagini al fine di testare la submission dell'algoritmo. Inoltre è presente anche un file **sample\_submission.csv** con un esempio di submission.

Come già detto i file train.csv e test.csv contengono i metadati per ogni immagine. Ogni metadato è etichettato con il valore 1 (Sì) o 0 (No):

- **Focus** - L'animale si staglia su uno sfondo ordinato, non troppo vicino/lontano;
- **Eyes** - Entrambi gli occhi sono rivolti in avanti o in avanti, con almeno 1 occhio/pupilla decentemente chiaro;
- **Face** - Viso discretamente chiaro, rivolto in avanti o vicino al davanti;
- **Near** - Singolo animale che occupa una porzione significativa della foto (circa oltre il 50% della larghezza o dell'altezza della foto);
- **Action** - Animale nel mezzo di un'azione (ad es. saltare);
- **Accessory**: accessorio/oggetto di accompagnamento fisico o digitale (es. giocattolo, adesivo digitale), esclusi collare e guinzaglio;

- **Group** - Più di 1 animale nella foto;
- **Collage** - Foto ritoccata digitalmente (cioè con cornice digitale, combinazione di più foto);
- **Human** - Umano nella foto;
- **Occlusion** - Oggetti indesiderati che bloccano una parte dell'animale (ad es. umano, gabbia o recinzione). Nota: Non tutti gli oggetti bloccanti sono considerati occlusione;
- **Info**: testo o etichette personalizzati (ad es. nome dell'animale, descrizione);
- **Blur** - Notevolmente sfocato o rumoroso, soprattutto per gli occhi e il viso dell'animale. Per le voci Sfocatura, la colonna "Occhi" è sempre impostata su 0.

## 2.1 Data Exploration

### TO-DO

In this section the available data sets must be presented. The term dataset refers to any type of information source, for example web services for geolocation fall into this category. In addition, all necessary data manipulation processes, such as cleaning and enrichment with external sources, must be presented and discussed.

## 3 The Methodological Approach

L'approccio proposto è basato su un modello multi-input single-output. La motivazione alla base di questo approccio è quella di sviluppare un sistema di apprendimento che sia in grado di sfruttare le informazioni utili da features miste di dati, poiché questi tipi di dati di solito richiedono un trattamento separato. A tal fine, ogni tipo di dato in ingresso viene elaborato e gestito in modo diverso e indipendente. L'architettura proposta è mostrata in figura 1

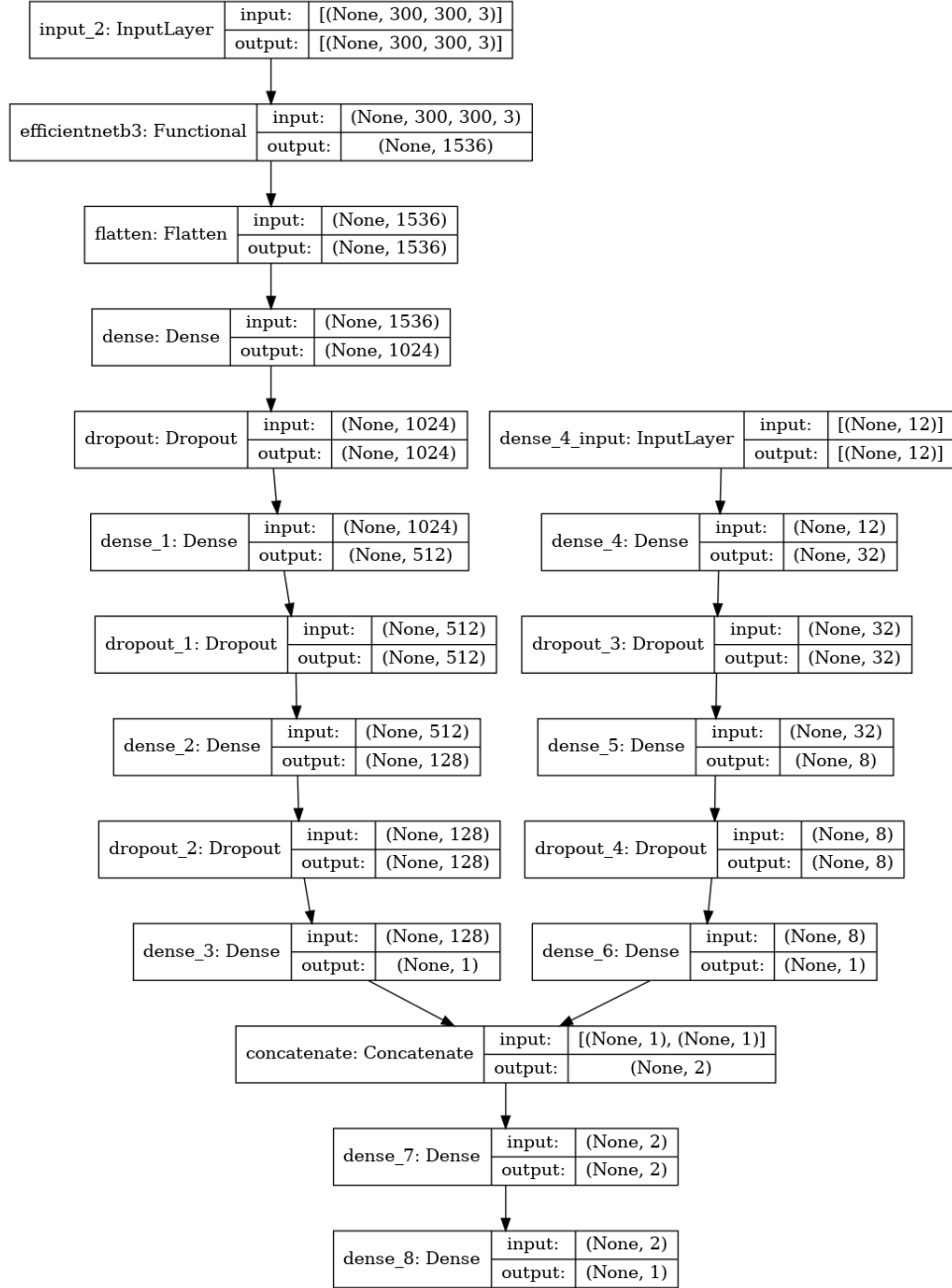


Figure 1: Multi-input Single-output Model

### 3.1

This is the central and most important section of the report. Its objective must be to show, with linearity and clarity, the steps that have led to the definition of a decision model. The description of the working hypotheses, confirmed or denied, can be found in this section together with the description of the subsequent refining processes of the models. Comparisons between different models (e.g. heuristics vs. optimal models) in terms of quality of solutions, their explainability and execution times are welcome.

Do not attempt to describe all the code in the system, and do not include large pieces of code in this section, use pseudo-code where necessary. Complete source code should be provided separately (in Appendixes, as separated material or as a link to an on-line repo). Instead pick out and describe just the pieces of code which, for example:

- are especially critical to the operation of the system;
- you feel might be of particular interest to the reader for some reason;
- illustrate a non-standard or innovative way of implementing an algorithm, data structure, etc..

You should also mention any unforeseen problems you encountered when implementing the system and how and to what extent you overcame them. Common problems are: difficulties involving existing software.

## 4 Results and Evaluation

The Results section is dedicated to presenting the actual results (i.e. measured and calculated quantities), not to discussing their meaning or interpretation. The results should be summarized using appropriate Tables and Figures (graphs or schematics). Every Figure and Table should have a legend that describes concisely what is contained or shown. Figure legends go below the figure, table legends above the table. Throughout the report, but especially in this section, pay attention to reporting numbers with an appropriate number of significant figures.

## 5 Discussion

The discussion section aims at interpreting the results in light of the project’s objectives. The most important goal of this section is to interpret the results so that the reader is informed of the insight or answers that the results provide. This section should also present an evaluation of the particular approach taken by the group. For example: Based on the results, how could the experimental procedure be improved? What additional, future work may be warranted? What recommendations can be drawn?

## 6 Conclusions

Conclusions should summarize the central points made in the Discussion section, reinforcing for the reader the value and implications of the work. If the results were not definitive, specific future work that may be needed can be (briefly) described. The conclusions should never contain “surprises”. Therefore, any conclusions should be based on observations and data already discussed. It is considered extremely bad form to introduce new data in the conclusions.

## References

The references section should contain complete citations following standard form. The references should be numbered and listed in the order they were cited in the body of the report. In the text of the report, a particular reference can be cited by using a numerical number in brackets as [1] that corresponds to its number in the reference list.  $\text{\LaTeX}$  provides several styles to format the references

## References

- [1] J. Lee, “Smart Factory Systems,” *Informatik-Spektrum*, vol. 38, no. 3, pp. 230–235, 2015.