BỘ GIÁO DỤC VÀ ĐÀO TẠO

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

**KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**

**🙞 🕮 🙜**



**MÁY HỌC NÂNG CAO**

**Đề tài:**

**ỨNG DỤNG DỰ ĐOÁN LOÀI ĐỘNG VẬT**

**Các sinh viên thực hiện:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tên:** | **Đỗ Trung Nguyên** | **Đỗ Thành Công** | **Nguyễn Hưng** |
| **Mã số:** | **B1709552** | **B1709526** | **B1709536** |
| **Khóa:** | **43** | **43** | **43** |

Cần Thơ

NHẬN XÉT CỦA GIẢNG VIÊN

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

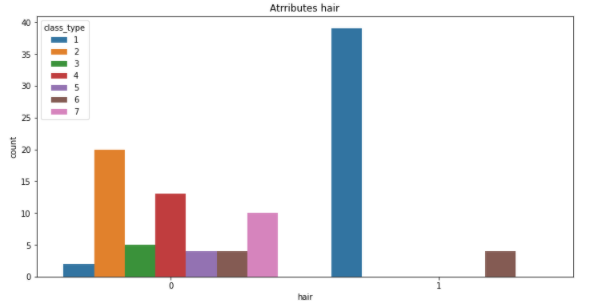
………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

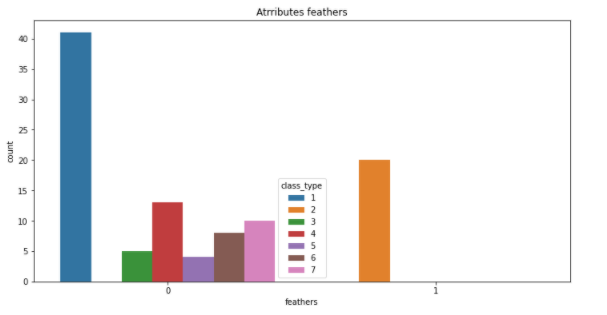
………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. **CÁC CÔNG VIỆC ĐÃ THỰC HIỆN**

* Thu thập dữ liệu từ nguồn Internet.
* Dataset Zoo gồm 2 file
* zoo.csv
* class.csv
* File zoo.csv gồm 17 trường thuộc tính chứa các đặc điểm đế phân loại loài động vật và 1 cột nhãn chứa tên các loài.
  + Thuộc tính
    - animal\_name ( Tên động vật )
    - hair ( Lông, tóc )
    - feathers ( Lông vũ )
    - eggs ( Sinh sản bằng trứng )
    - milk (Có sữa )
    - airborne ( Môi trường trên không )
    - aquatic ( Môi trường dưới nước )
    - predator ( Động vật ăn thịt )
    - toothed ( Răng )
    - backbone ( Xương sống )
    - breathes ( Hô hấp )
    - venomous ( Có nộc độc )
    - fins ( Vây )
    - legs ( Chân )
    - tail ( Đuôi )
    - domestic ( Bầy đàn )
    - Catsize.
  + Nhãn
    - class\_type
      * 1 : Động vật có vú
      * 2 : Chim
      * 3 : Bò sát
      * 4: Cá
      * 5 : Lưỡng cư
      * 6 : Bọ
      * 7 : Không xương sống
* Vẽ một số biểu đồ tương quan giữa các thuộc tính với nhãn trong tập dữ liệu.
* Lông (hair)

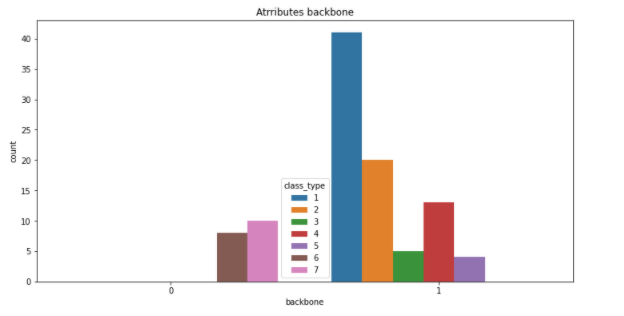


Các loài nằm trong khung tỉ lệ 0 và 1, da số động vật thuộc ‘động vật có vú’ thường có bộ lông, tóc và một số ít không có lông, tóc. Cũng như lớp 6 ‘bọ’ thì không có lông.

* Lông vũ (feathers)

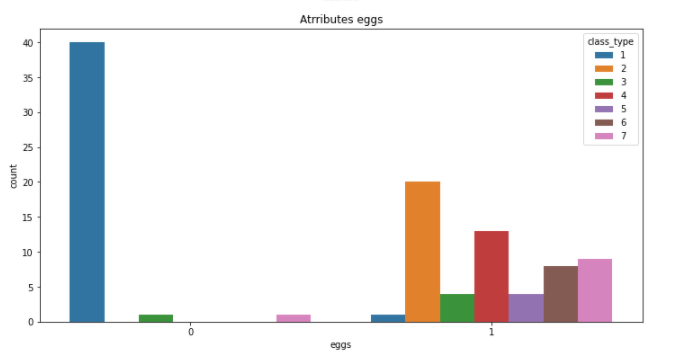
Theo biểu đồ thì loài 2’chim’ thì 100% là có lông vũ. Và các loài khác thì rải rác từ 4, 5, 6, 7.

* Xương sống (backbone)



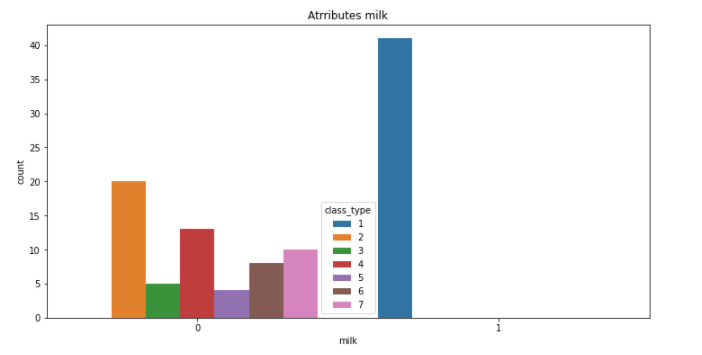
Về thuộc tính xương sống thì đa phần các loài trong nhãn đều có xương sống, ngoại trừ loài 6 ‘bọ’ và loài ‘7’ không xương sống .

* Trứng (eggs)



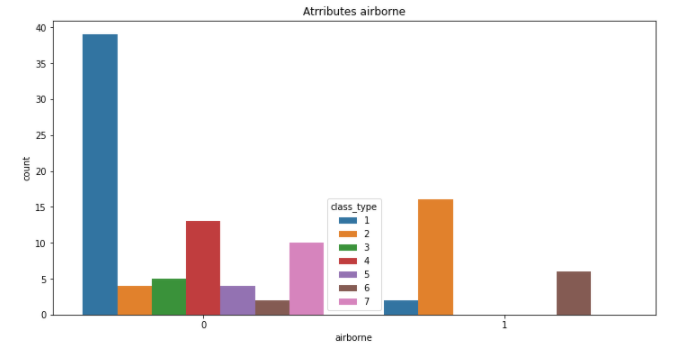
Đa phần các loài thuộc hệ ‘2’ chim, ‘3’ bò sát và ‘4’ cá, thường đẻ trứng và một ít loài lưỡng cư.

* Sữa (milk)



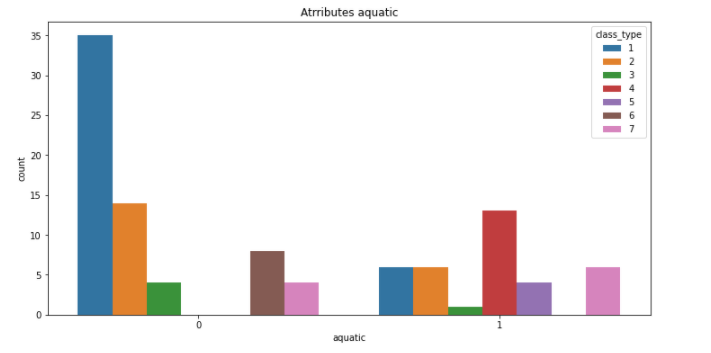
Biểu đồ chỉ rõ động vật có vú ‘1’ là có sữa.

* Môi trường trên không (airborne)



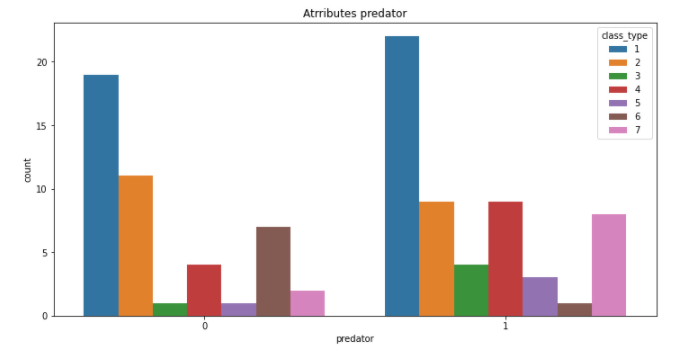
Hệ chim ‘2’ là hệ tỉ lệ cao sống ở môi trường trên không và một ít loài động vật có vú.

* Môi trường dưới nước (aquatic)



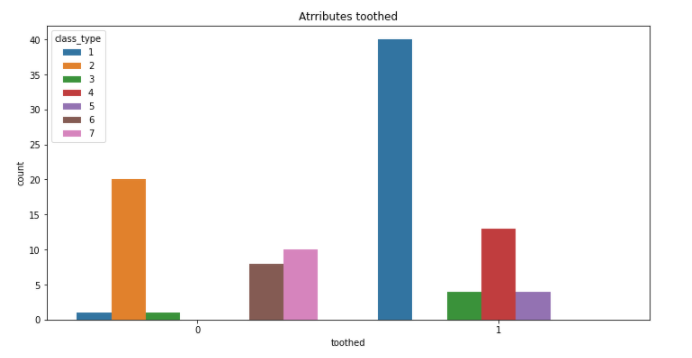
Một số loài trong các hệ ‘1’ động vật có vú, chim ‘2’, bò sát ‘3’và phần lớn loài hệ cá ‘4’ sinh sống và hoạt động được ở môi trường trong nước.

* Động vật ăn thịt (predator)



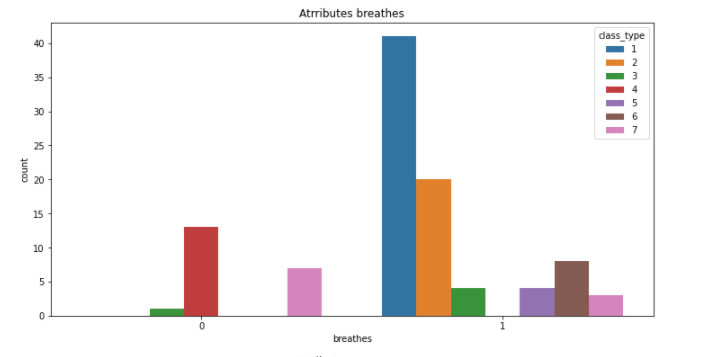
Các loài thuộc hệ ‘1’ động vật có vú, ‘2’ chim , bò sát ‘3’, bò sát ‘4’ và một ít loài hệ ‘5’ ‘6’ thường ăn thịt và cũng có một số loài không ăn thịt.

* Có răng (toothed)

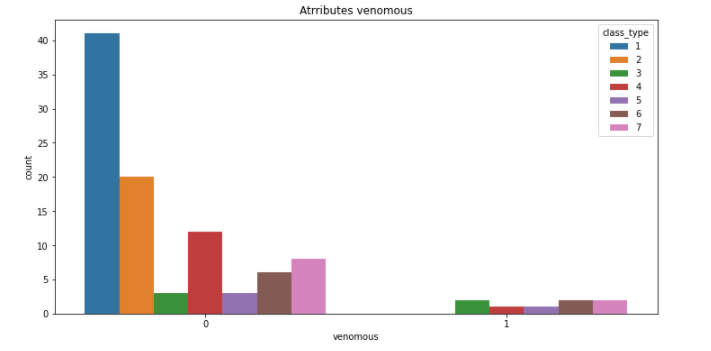


Hê ‘1’ động vật có vú thường có răng và một ít loài thuộc hệ ‘3’ bò sát và ‘4’ cá có răng.

* ….(breathes)

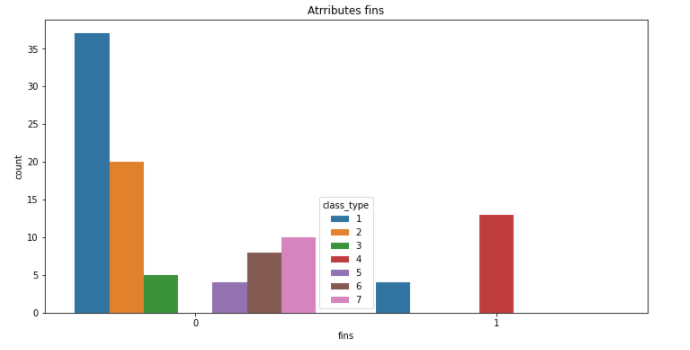


* Có nộc độc (venomous)



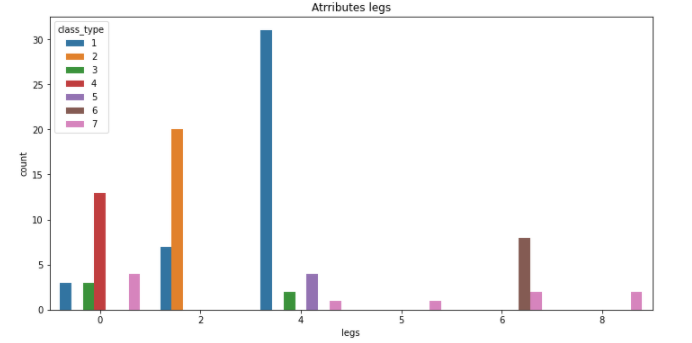
Một ít loài hệ cá ‘3’, bò sát ‘4’, lưỡng cư ‘5’, bọ ‘6’ và không xương sống ‘7’ có nộc độc trong cơ thể. Đa phần là không có.

* Vây (fins)



Cá loài thuộc hệ cá ‘4’ là có vây, các hệ khác hoàn toàn không có.

* Chân (legs)

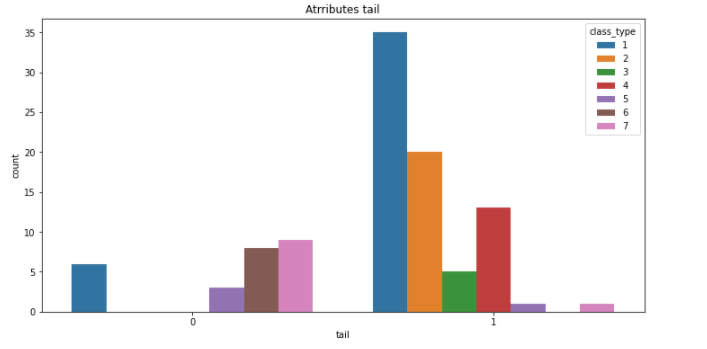


Các hệ rải rác từ 0,2,4,5,6,8 chân tùy hệ. Phần lớn hệ động vật có vú ‘1’ thuộc 4 chân, chim ‘2’ thuộc 2 chân. Và các loài trong các hệ rải rác.

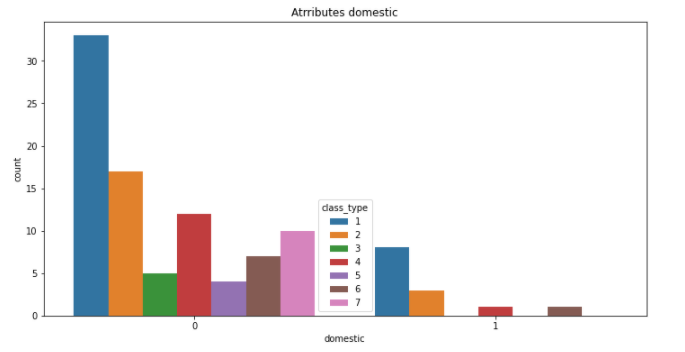
* Đuôi (tail)

Các loài trong các hệ trong nằm trong thuộc tính đuôi gồm:

‘1’ động vật có vú, ‘2’ chim , ‘3’ bò sát’ , ‘4’ cá.

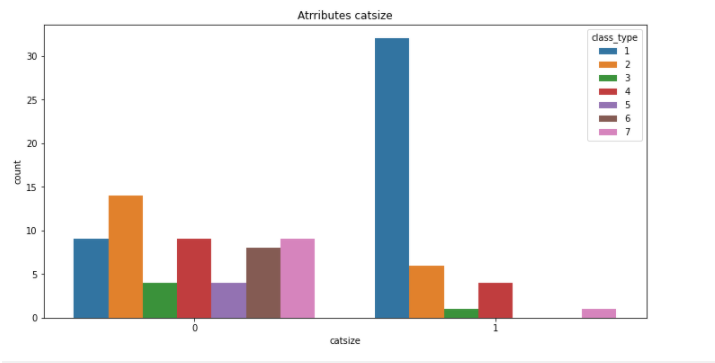


* Bầy đàn (domestic)

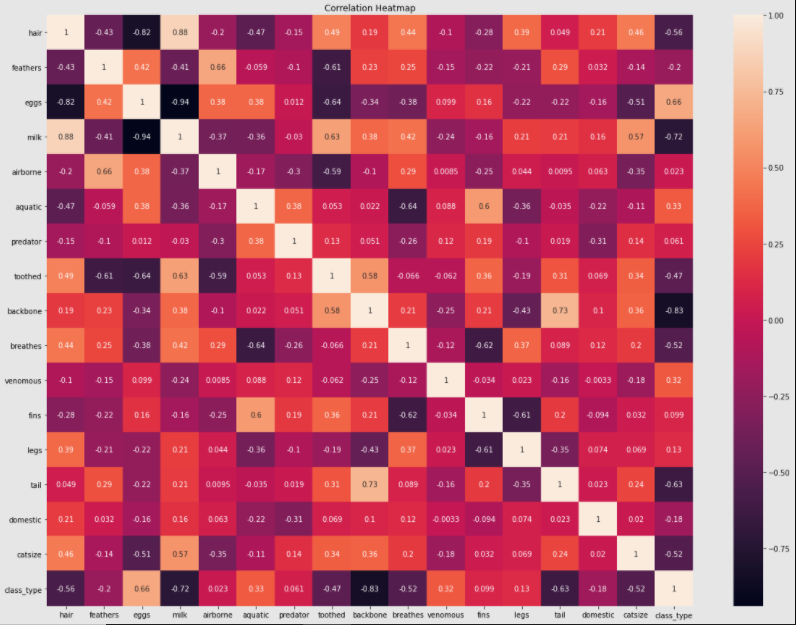


Theo biểu đồ nhận thấy chỉ một số loài trong hệ ‘1’động vật có vú, ‘2’ chim và ‘4’ cá sống theo bầy.

* (catsize)



* Vẽ biểu đồ nhiệt thể hiện giá trị tương quan giữa các thuộc tính trong hệ với nhau.



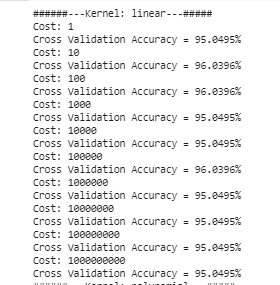
* Sau khi biết được sự tương quan giữa thuộc tính với nhãn. Sử dụng thuật toán svm để huấn luyện mô hình.
* Sử dụng các kerner khác nhau để so sánh kết quả.
* Với mỗi kerner sử dụng các tham số c từ 1,10,100,…,1000000000. Để so sánh kết quả.

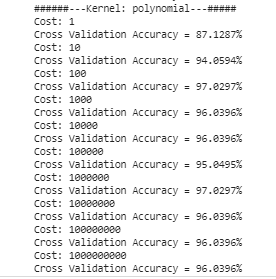
0 : linear

1 : polynomial

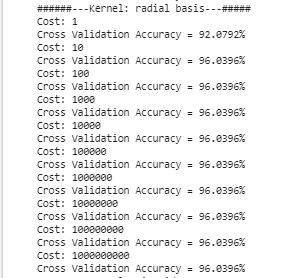
2 : radial basis

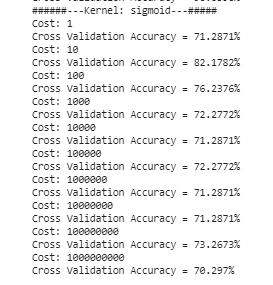
3 : sigmoid

* Kết quả train từ các kernel.
  + Linear
* Polynomial



* Radial basis

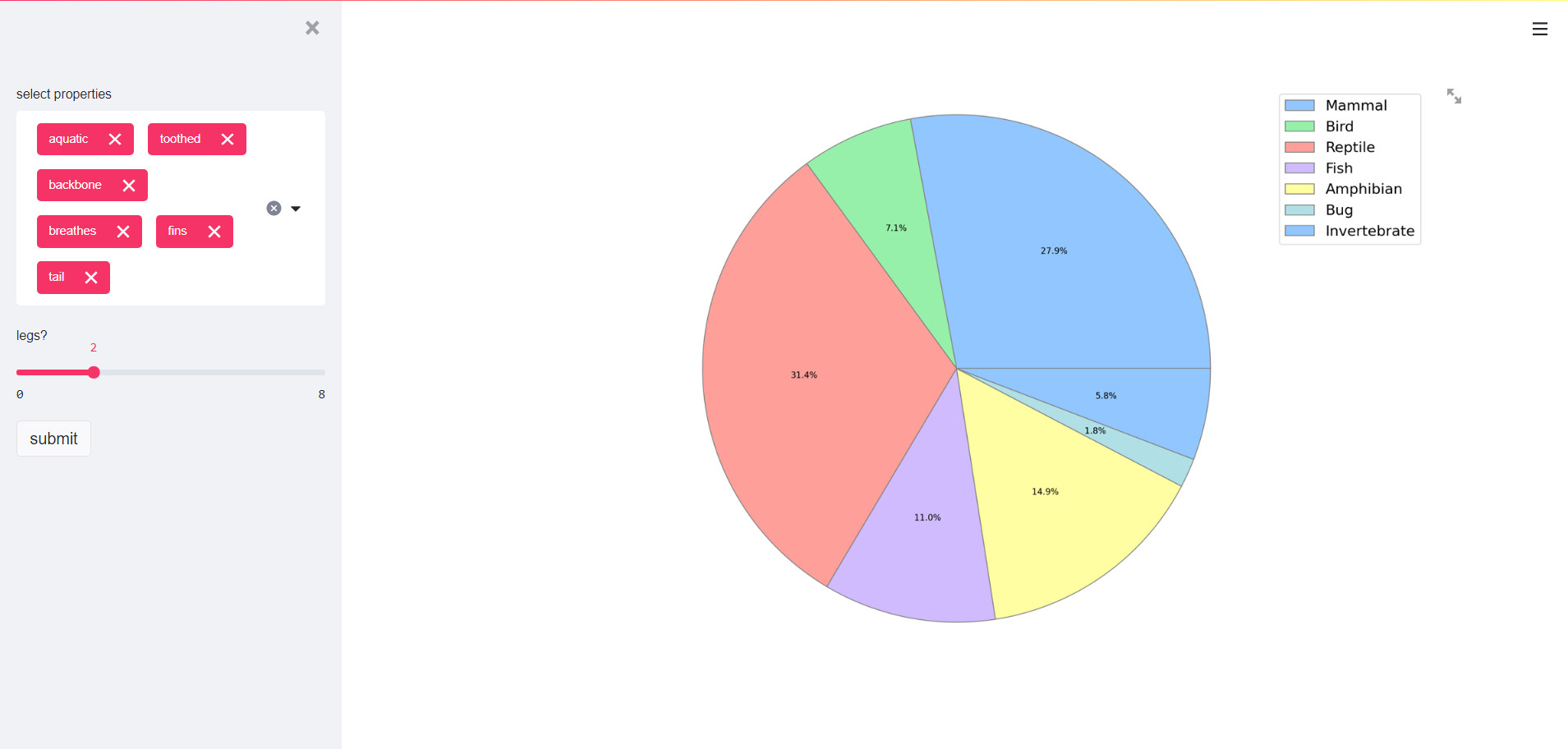


* Sigmoid
* Sau khi train mô hình sử dụng tham số c từ 1 – 1000000000. Ta được biểu đồ đánh từ sau c từ 1 đến 1000.
* Từ biểu đồ trên ta nhận thấy được sự ổn định của kernel ‘linear’ khi áp dụng cho mô hình.
  + Sử dụng kernel linear cho tạo mô hình và đánh giá.
* Xây dựng giao diện sử dụng

Tạo giao diện web đơn giản bằng…

Hình ảnh giao diện dự đoán cho :

Gozila = [aquatic, toothed, backbone,breathes , fins, tail] + 2 legs



Kết quả cho thấy ….

1. **BẢNG PHÂN CÔNG**

|  |  |  |  |
| --- | --- | --- | --- |
| Công việc | Thành viên | Kết quả | Ghi chú |
| Thu thập dữ liệu |  |  |  |
| Phân tích dữ liệu |  |  |  |
| Vẽ các sơ đồ tương quan |  |  |  |
| Sử dụng libsvm |  |  |  |
| Đánh giá |  |  |  |
| Xây dựng giao diện ứng dụng |  |  |  |
| … |  |  |  |