Proposal Documentation

Release 1.0

Team EagleBytes

Oct 05, 2017

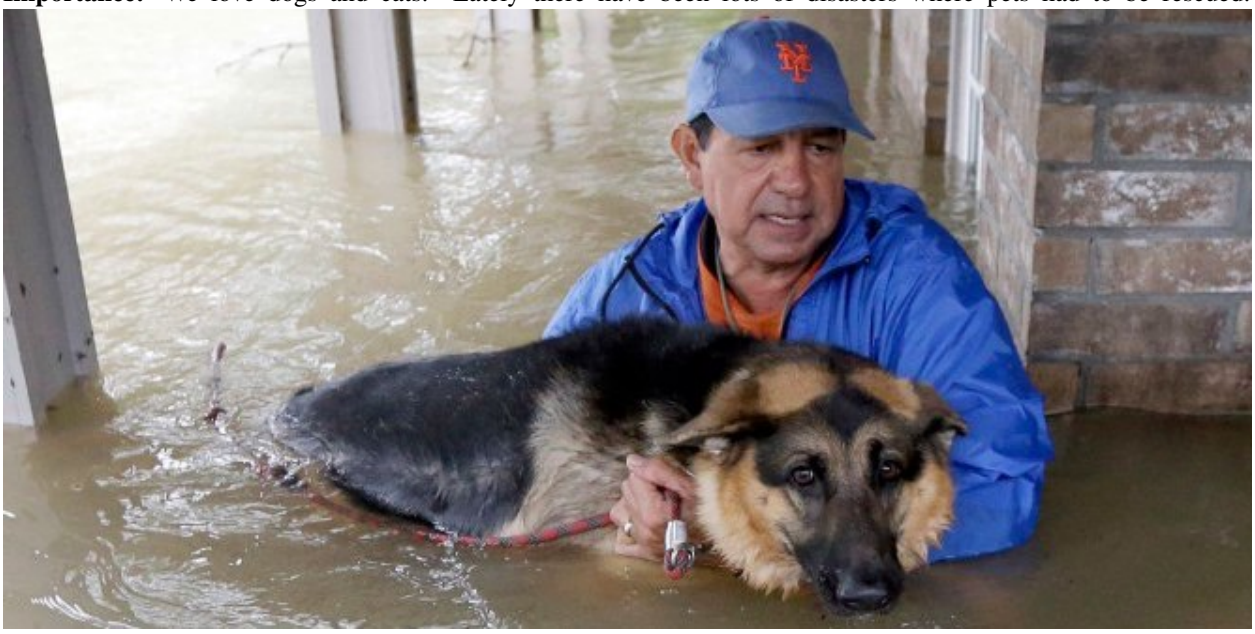
SUBSECTIONS:

Team ID: PESSF118, Pinon Elementary School, Santa Fe

Science Area: Statistics

Hypothesis: We can use data from social media to tell how likely it is that pets are present at a given street address.

Importance: We love dogs and cats. Lately there have been lots of disasters where pets had to be rescued:



In disasters, pets often get separated from their humans and other people have to help, like firemen, policemen, animal rescue volunteers, or just good neighbors. It can be hard for rescuers to know where pets need to be rescued and where they don't.

People don't always bother to put stickers on their houses that list their pets, but lots of people post pics online of their dogs and cats. Most pics of pets are taken at home. Sometimes you can tell where a pic was taken, especially if the post has a location tag. We think if you combine these things you might be able to say whether there is a pet at an address needing to be rescued in a disaster.

When we explained our idea in an e-mail to Jay Mitchell, Cabinet Secretary for the New Mexico Department of Homeland Security and Emergency Management, he responded:

```
I have to lead off with "wow", that's a very hefty and commendable project.
Pet care, pet evacuations and companion animal support is a very important
and at times, a difficult task during emergencies and disasters.
```

Another person who agreed to help us is David Silver, Emergency Operations Manager for the City of Santa Fe. They will put us in touch with firemen, people who monitor social media in disasters, and volunteer pet rescuers.

Even if our hypothesis is wrong, we might still help pets in disasters by inventing new ways to use social media and telling people about it.

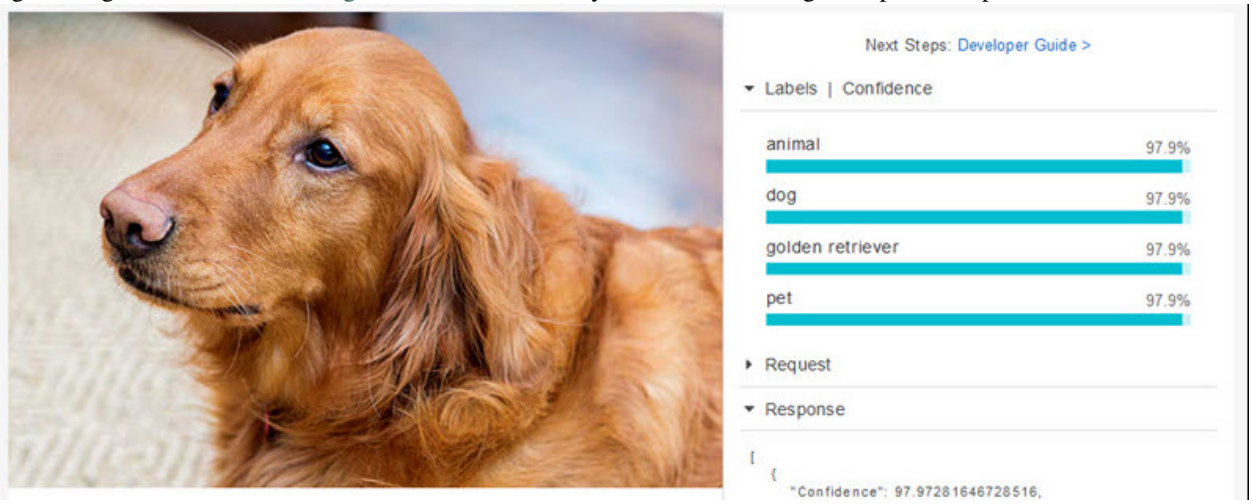
CHAPTER

ONE

APPROACH

We plan to use Python for this project because there are examples in Python of how to do each of the steps. We plan to do visualization on the web in Python with plot.ly's Dash because it is interactive and easy to use.

We plan to use Amazon Web Services (AWS) for this project because it has good support for data supercomputing projects like this one, with ways to scale our project to hundreds of servers if we need to. Also AWS gives us some free use through the AWS Educate program, and it has an image recogition service, Rekognize, that is already trained to recognize pets in pictures like this one:



We already have a web server on AWS running for our project at http://EagleBytes.org that we will use to do calculations and present results. We will keep our code in a GitHub repositories accessible from https://github.com/EagleBytes2017 .

We have to do more research before we can say which social media sources we will use for this project, but we are considering Facebook, Twitter, Instagram, and Tumblr.

To test our hypothesis, we plan to compare results of our calculations with results from surveys we do in our own neighborhoods. We plan to get help from our parents in going door to door and asking questions from our neighbors about social media use and pets.

1.1 Weaknesses

We don't think any Challenge team has done a data supercomputing project before; "Data Science" isn't on the list of scientific areas. The problems we face are going to be different than those other teams have faced.

We are a young team, and we won't have the programming skills of older kids.

We think our biggest problem is going to be access to data. Some sites have but don't allow access to data. Other sites make it easy to access data, but they may charge money. We plan to get advice from disaster social media volunteers, then we will write social media companies to request access. If we can't get the access we need, we may have to focus on simulating data by posting with hashtags we tell people to use.

Another problem is that when we want to scale up our calculations, we can't use the computers at LANL because they aren't set up for data supercomputing. On AWS, we may be limited by the amount of service we can afford.

CHAPTER

TWO

TEAM

2.1 Team Members

Team members' names are withheld in accordance with COPPA rules. Their secret hacker identities are:

- DarkDJ
- Superskatingdog
- soccerchamp
- spursqueen
- volleyballqueen
- Warmachine

2.2 Sponsoring Teacher

Delara Sharma

2.3 Project Mentor

Joel Berendzen (GenerisBio, LLC)