

# Combinatorial Data Augmentation on Satellite Images

- Summer Internship at Radix -

## About Radix

Radix is an Artificial Intelligence solutions provider based in Brussels. Together with our clients, we deliver AI solutions that have a positive impact on the world.

Our mission is to help our clients grow and to improve people's lives. To deliver on our promise, we focus on the combination of three elements:

- Business: creating high value for our clients, their employees and society as a whole
- People: delivering technology that can truly assist people, not replace them
- AI: automating intelligence to achieve the highest impact

Founded in January 2018, Radix is now a team of 25 people that continues to grow. We work with ambitious clients like GSK, Brussels Airport, Atlas Copco, Flanders Investment and Trade, VDAB, Belga, Macadam, House of HR and more.

Visit our website <https://radix.ai> for more information.



## The internship

As of this year, Radix will host several summer internships. Such an internship shows you, among other things, how your learning material is applied in the field and what the best practices are when working on a project, this while being supervised by field experts. Aside from that, it allows you to get a glance of how a growing startup operates from the inside and how other fields such as *Marketing* and *Sales* contribute to this.

### About the project

Data augmentation helps increase the size of your training dataset which is often required in practical applications of Machine Learning. Standard approaches are to apply scaling, rotations, cropping, etc. In this internship, you will try out a creative way of data augmentation based on combining and aggregating data points. The key idea is that while for a dataset of  $N$  samples, this  $N$  might be too small to train the model sufficiently, while this wouldn't be the case when two or three variations of each of the  $N$  samples were added to the dataset. You can apply this idea to a crop yield prediction dataset (satellite data) or a dataset of your choice as long as the technique applies to the dataset.

### What you will learn

In this internship, you will learn about the importance of dataset size, what this implies in practice, and the different solution to circumvent this problem. You'll learn to do a ML experiment in a real-life setting and think about what its impact could be. Lastly, you will learn to work with imagery data.

## Your profile

- You are in your third Bachelor or first Master in Computer Science or a related field
- You are fluent in English, both written and verbally
- You have strong analytical skills and are familiar with the basics of machine learning
- You know your way around in Python
- Familiarity with statistical Python tools such as Spacy, TensorFlow, or Torch is a plus
- You are eager to learn and are open to be challenged