Internship:

Manifold hypothesis testing for common datasets

I.	CONTEXT

The Global Research Program is a virtual international experience that supports research and creative projects in UC Chile and international partner universities within the Global Internships for Research and Practice-Based Learning.

II.	INTERNSHIP DETAILS

Name	Manifold hypothesis testing for common datasets
Description	A dataset (i.e. a point cloud in R^d, cases of interest being those with $d>10^2$) satisfies a "manifold hypothesis" if it is well-approximated by a (low dimensional) submanifold. In theory, if the manifold hypothesis is satisfied, this allows to effectively lower data dimensionality and improve the efficiency of several classes of algorithms, including Deep Learning algorithms. Efficiently testing whether this hypothesis is satisfied (or even just determining good dimension bounds for the approximating manifold) is itself a very challenging task, mainly due to the high dimensionality of datasets of interest. The proposed project aims at implementing and experimenting with new algorithms for effective testing of the manifold hypothesis based on multi-scale analysis of the data. We are developing a new Geometric Measure Theory theorem which ensures the existence of the approximating manifold under low covariance bounds for the restriction of the dataset to a collection of balls, and the project would consist of creating new algorithms for implementing this method, as well as optimizing the method based on the outcome of experiments with real datasets.
Internship Language	English, Spanish, French, Italian
Places Available	2
Application Period	January 27, 2022 - February 28, 2023.
The desired academic level of the students	Senior-year undergraduate students, master's students.
Desired program/discipline of the students	Mathematics, Computer and Information Sciences, Electrical Engineering, Electronics, and Computer Science.

Specific skills that a student should have for the project	Required: - Advanced programming experience (for implementing the multiscale algorithms and understanding/presenting outcomes of tests on real data) - Familiarity with basic mathematical analysis and probability notions (for understanding the terminology and so that it is possible to learn, during the project, the underlying ideas of why the algorithms should work) Suggested: - Familiarity with basic Deep Neural Network architectures and their behavior (we may use DNN classification accuracy as a test for success of our manifold detection algorithms) - Knowledge of any of the following is not indispensable, but will allow to "get a better taste" of the ideas we use: geometry of curves and surfaces, basic differential geometry (tangent spaces and measures of intrinsic/extrinsic curvature), concentration of measure
Duration for the research internship	12 weeks
Desired dates for the research internship	From: 4/1/2023 (April 1st, 2023)To: 6/30/2023 (June 30th, 2023)
Benefits	● An acceptance letter to participate in a virtual internship at UC Chile. ● A participation certificate issued by the professor as long as all tasks assigned have been completed. ● Access to our virtual library catalogue. ● All students will have access to a virtual workshop - preparation package before their research internship.

Please send all the documents via email to cristian.diaz@uc.cl, Director for Global Mobility at UC | Chile.