Mathematical and statistical models for Big Data in the agri-food sector

Topic for a thesis in Applied Statistics-Mathematics through a Cifre contract with NeoVia, supervised by the LMBA and SEE-d

Context
The NeoVia company (www.invivo-nsa.com), a subsidiary of the InVivo Group (www.invivo-group.com) is currently conducting a project regarding a connected farm. In this context, large heterogeneous data streams are received from the sensors embedded in the animals and in their environment.
The data are intended to show the state of the livestock and help manage it accurately.
NeoVia intends to develop innovating Machine Learning methods that are specifically adapted to the agri-food sector.

Topic
The purpose of this applied mathematical and/or statistical thesis is to contribute to constructing a tool, making synthetically the links between what the animals endure (food, genetics, etc.), their state (rumination, gestation, calving, health status, etc.), and the data sent by the sensors.
Once these links are established, the tool must use the data received from the sensors and the characteristic features of an animal to predict the animal's state and, if need be, to prescribe a response to help improve its state of health.

This approach, which involves linking input data to output data through an operator adjusted on the basis of a learning base of historical data, is common in the fields of Machine Learning and Big Data.
The distinctive feature of the topic is the method that will be implemented to construct this operator.

Living organisms are the scene of numerous processes such as flow, distribution, delay, accumulation, etc., and each of these processes is converted into mathematical differential and integral mathematical equations that may be defined in the same domain.
The tool to be constructed will be based upon a mathematical model showing how coupling these types of processes can convert input data (food, genetics, etc.) into output data (similar to what the sensors capture) by taking into account the influence of the animal's state and how that, in turn, is influenced by what the animal endures. Weighting the various processes in the coupling model and positioning the input and output data in the model will not be imposed; these will come from a learning base.

Translating the mathematical model into a technological tool will require numerical analysis methods.

The tool will be a new component - oriented to the analysis of bio-based Big Data - of the set of algorithms of Machine Learning.

**Doctoral candidate**

Holding a Master's or a diploma in statistics or applied mathematical engineering, you enjoy working with and are competent in mathematical and statistical modeling and in data analysis. You know R software, database management and NoSQL technologies; you have knowledge of numerical analysis, and you can program. You enjoy developing tools to use data to bring concrete and relevant responses to issues. You are interested in developing innovative methods based upon high-level mathematical and statistical modeling for a dynamic and pioneering company. Commitment and hard work are your basic values. You wish to write your doctoral thesis within a multidisciplinary team including engineers, doctors, company managers, and academics, at the interface between the business and research worlds.

If you see yourself in the description above, NeoVia wishes to recruit you to conduct research under a Cifre contract.

Please send your application (CV, motivation letter and URL) by email to:
- Vincent Sincholle ([vsincholle@neovia-group.com](mailto:vsincholle@neovia-group.com)), and
- Emmanuel Frénod ([emmanuel.frenod@univ-ubs.fr](mailto:emmanuel.frenod@univ-ubs.fr))

Supervision will be jointly conducted by
- Vincent Sincholle's team at Neovia in Saint-Nolff;
- the applied mathematics and statistics team of the Laboratory of Mathematics of Bretagne-Atlantique (LMBA – UMR6205) in Vannes; and
- See-d ([www.see-d.fr](http://www.see-d.fr)) in Vannes.