the produced results. We need to develop techniques that will produce answers much earlier than running the Neighbor similarity search, and k-Nearest Neighbor classification. In all cases, we need to develop techniques that will produce answers much earlier than running the full algorithm, and at the same time provide probabilistic guarantees on the quality of the produced results.

References:


Mission confiée

We have developed a probabilistic learning-based method that provides probabilistic time and distance guarantees for progressive 1-Nearest Neighbor (1-NN) query answering. Our approach learns from a small set of queries and builds statistical models that assess how far we are from exact answers. We can then communicate this information to users, and users can decide to terminate a progressive search to reduce waiting times.

The goal of the internship is to develop new techniques that extend this approach to more complex data-series analysis tasks, for example, to the problems of k-Nearest Neighbor similarity search, and k-Nearest Neighbor classification. In all cases, we need to develop techniques that will produce answers much earlier than running the full algorithm, and at the same time provide probabilistic guarantees on the quality of the produced results.
Compétences
The candidate is expected to have a Computer Science degree (M1/M2 level), knowledge of statistical analysis methods, and experience with C/C++ and R programming.

Avantages
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage