**ENS / IISER collaboration**

***Internship subject form***

To be sent back by January 31st, 2020

|  |  |
| --- | --- |
| Name of the institution | Ecole normale supérieure Paris-Saclay |
| Name of the host laboratory | LaboratoireSpécification et Vérification |
| Website of the host laboratory | www.lsv.fr |
| Research group | Inriateam Mexico |
| Internship number | COMPUTER SC |
| Internship subject (title) | Specification and Verification  of Properties of Neural Networks |
| Prerequisites | Some knowledge about neural networks  and/or verification techniques |
| Internship proposal: description and expected training outcomes (15 lines max.) | With the development of machine learning and its daily applications, gainingconfidence in the systems produced by such techniques has become a criticalissue. A first problem consists in formalizing what is expected from the systems. Such requirements may be either generic or specific to the task to be achieved.For instance, adversarial robustness is a generic property. It measures howmuch information is needed by an attacker to \falsify" the answer of a classifyingsystem. On the other hand, assume the system proposes actions to be performedin the presence of an intruder, a specific property would be that there is no actionto be proposed when no intruder is detected.In the internship, we will focus on neural networks since this is the mostwidely used and moreover it presents similar features to hybrid systems lettingthe possibility to adapt effectivetechniques from this domain.Thus the goals of this internship are twofold: (1) Specifying a language or a logic that can express the main properties expected to be satised by neural networks; (2) Identifying specificities of formula related to these properties in order to design new exact and/or approximate algorithms for verifying these properties. |