

---

# Machine Learning and application to text mining

Ana Karina Fermin<sup>\*†1</sup>

<sup>1</sup>Université Paris Nanterre (UPN) – Université Paris Nanterre – France

## Résumé

Au cours de ce tutoriel, nous analyserons de manière automatique des textes de Victor Hugo. À l'aide de ce corpus, nous verrons des exemples des tâches les plus importantes en "text mining". Nous parlerons de lecture, nettoyage, prétraitement des textes, de leur classification non supervisée et supervisée. Chaque technique sera précédée d'une brève introduction permettant d'en donner les grands principes.

- **Ana Karina FERMIN** est maître de conférences à l'université Paris Ouest Nanterre La Défense. Elle effectue sa recherche dans l'équipe Modal'X et enseigne les statistiques dans différents départements de son université. Ses thèmes de recherche vont des problèmes inverses à l'apprentissage statistique avec des applications dans des domaines variés.

---

\*Intervenant

†Auteur correspondant: [aferminrodriguez@u-paris10.fr](mailto:aferminrodriguez@u-paris10.fr)