

BAHAREH AFSHINPOUR

Jeune chercheuse (post-doctorat)

Mariée, 2 enfants

Université Grenoble Alpes
Laboratoire d'Informatique de Grenoble
Bâtiment IMAG
150 place du Torrent
38401 Saint-Martin d'Hères

Email : bahareh.afshinpour@univ-grenoble-alpes.fr
Web : www.afshinpour.com
LinkedIn : [linkedin.com/in/bahareh-afshinpour](https://www.linkedin.com/in/bahareh-afshinpour)
Adresse personnelle : France.

Études - Diplômes

- 2023 Diplôme de docteur**, spécialité informatique (Université Grenoble Alpes, Grenoble, France)
Titre : *Exploitation de journaux logiciels avec des techniques d'apprentissage automatique*
Directeurs de thèse : Roland Groz, Massih-Reza Amini, Professeurs des universités, Université Grenoble Alpes.
Membres du jury : Présidente : Oum-El-Kheir Aktouf, Examineur : Arnaud Gotlieb, Rapporteurs : Neil Walkinshaw et Franz Wotawa.
- 2010 Diplôme de master en sciences, domaine de l'ingénierie informatique et logicielle.** (Université Payam-e-Noor, Téhéran, Iran).
Mémoire de Master : *Nouvelle méthode d'optimisation du classement des résultats de recherche sur le Web sémantique basée sur des ontologies*. Directeur : Hassan Khotanlou, Professeur de génie informatique, Université Bu-Ali Sina, Hamedan, Iran.
Mention : Excellent.
- 2005 Licence en génie logiciel (Diplôme d'ingénieur en génie logiciel)** à l'université Payame-Noor, Hamedan, Iran.

Parcours Professionnel

- 05/2024 - 09/2025** Post-doc dans le cadre du projet "*Conception d'un module d'analyse multimodale des sentiments, audio et vidéo, pour le développement d'un agent conversationnel*" dans l'équipe apprentissage machine (APTİKAL) du Laboratoire Informatique de Grenoble (LIG) à l'Université Grenoble Alpes, Grenoble, France.
- 2023-24** Ingénieure de recherche, dans le cadre du projet EFELIA-MIAI¹, à l'Université Grenoble Alpes, Grenoble, France.
Conception d'un programme de master en IA (cours/TD/TP) niveau M1 & M2 pour des étudiants de Sciences-Po, Grenoble
- 2022-23** Attachée temporaire d'enseignement et de recherche (A.T.E.R.), Informatique, à l'UFR, IM2AG, l'Université Grenoble Alpes, Grenoble, France.
- 03/2019-09/2019** Ingénieure, au laboratoire TIMC, Université Grenoble Alpes, Grenoble, France.
- 2017-2019** Chercheuse, au Laboratoire LBB (Biologie des systèmes et bioinformatique), Université de Téhéran, Iran. (J'avais entamé une thèse de doctorat en bioinformatique en Iran, mais j'ai dû l'arrêter suite au déménagement de ma famille en France).
- 2010-2019** Ingénieure de formation, au département de génie informatique, université Payam-e-Noor, Téhéran, Iran. Enseignement de cours de licence (L1, L2 et L3).
Planification et coordination de l'horaire de tous les cours de L1, L2 et L3 en contactant leurs responsables de cours et leurs enseignants.

Enseignement

J'aime enseigner et j'ai toujours été intéressée par les pédagogies actives ; ainsi, j'ai suivi de nombreuses formations pour me mettre à jour sur les outils informatiques et les nouveaux langages de programmation (tels que Python avancé). J'ai eu la chance d'enseigner l'informatique et les mathématiques discrètes à tous les niveaux universitaires en licence et master.

1. <https://miai.univ-grenoble-alpes.fr/education/efelia-miai-1400610.kjsp>

Conception de cours

En 2023, j'ai occupé un poste d'ingénieure de recherche dans le cadre du projet Compétences et Métiers d'avenir France 2030, EFELIA-MIAI. Mes responsabilités incluaient :

- La conception de cours en intelligence artificielle (IA) pour les étudiants non informaticiens de master de Sciences Po Grenoble (IEP), dont le responsable est Olivier Le Van Truoc, professeur associé à Sciences Po Grenoble (olivier.le-van-truoc@iepg.fr).
- L'élaboration d'un guide pédagogique à destination des enseignants-chercheurs pour les aider à dispenser ces formations.

J'ai également été impliquée dans la réflexion et le développement de projets pédagogiques visant à sensibiliser et à former les étudiants du Master Progis (Data intelligence et études Opinion, marketing, médias) à l'IA et à ses enjeux professionnels et sociétaux. Pendant l'année scolaire, les étudiants de Master Progis sont en alternance et sont formés d'environ un cours par mois.

Dans ce cadre, j'ai créé des maquettes de cours détaillées pour ces étudiants. Ces cours, structurés sur une période de 12 semaines, sont conçus pour offrir une formation complète et progressive IA. Ils comprennent des cours magistraux, des travaux dirigés (TD) et des travaux pratiques (TP), pour une acquisition des compétences tant théoriques que pratiques dans l'objectif d'initier les étudiants à l'algorithmique et à l'IA de manière approfondie. J'ai l'opportunité de les présenter aux étudiants durant l'année scolaire 2024-2025.

Le programme que j'ai mis en place, débute par une introduction aux concepts de base de l'apprentissage machine, en explorant les différentes techniques et algorithmes fondamentaux. Au fur et à mesure de l'avancement du cours, les notions deviennent progressivement plus complexes, telles que les perceptrons multi-couches (PMC), les réseaux de neurones convolutifs (CNN), et les réseaux de neurones récurrents (RNN). Les travaux pratiques consistent à la mise en application des concepts théoriques à travers des projets concrets, afin d'offrir aux étudiants une expérience pratique. En outre, le cours intègre des discussions sur les implications éthiques et sociétales de l'IA, afin de sensibiliser les étudiants aux enjeux plus larges de cette technologie. Des études de cas et des projets collaboratifs sont également inclus pour encourager le travail en équipe et la résolution de problèmes complexes.

En somme, ces maquettes de cours ont visé à fournir aux étudiants une compréhension solide et pratique de l'IA, les préparant ainsi à relever les défis professionnels et à contribuer de manière éclairée à la société algorithmique.

Certains éléments du matériel pédagogique et des fichiers sont disponibles sur ma page personnelle : <https://afshinpour.com/courses.html>.

Activité d'enseignement

Depuis 2010, j'ai accumulé une riche expérience dans l'enseignement à différents niveaux académiques, totalisant 1771 heures de cours.

En Licence, j'ai enseigné diverses matières en informatique et en sciences de la vie. En L1, j'ai dispensé 819 heures de cours, couvrant des sujets tels que la programmation en C++ (48 heures par an), les mathématiques discrètes (24 heures par an), et l'informatique pour les sciences de la vie (33 heures par an). La responsable informatique de cette matière est Lydie du Bousquet, professeure à l'université Grenoble-Alpes, lydie.du-bousquet@univ-grenoble-alpes.fr

En L2, j'ai enseigné l'algorithmique (36 heures par an), totalisant 360 heures. Ces cours étaient basés sur l'ouvrage édité que j'avais rédigé en 2015 [2].

En L3, j'ai donné des cours en ingénierie logicielle et en conception de sites web (24 heures par an pour chaque matière), accumulant 480 heures.

En Master, mon enseignement s'est concentré sur des sujets plus spécialisés. En M1, j'ai enseigné les fondements des bases de données (41.25 heures) et l'introduction aux bases de données (36.63 heures), totalisant 77 heures, pour des étudiants du parcours Master of Science in Informatics at Grenoble (MoSIG) et Master of Applied Mathematics (AM). Le responsable de ces cours est Cyril Labbé, professeur à l'université Grenoble-Alpes, cyril.labbe@imag.fr.

En M2, j'ai dispensé en 2024 un cours sur l'apprentissage automatique dans le cadre du Master Progis (35 heures) en utilisant les supports de cours que j'ai présentés plus haut.

Mon engagement pédagogique s'étend sur plusieurs années. Cette expérience diversifiée m'a permis de développer des compétences solides en pédagogie, en adaptation des contenus aux différents niveaux d'études, et en interaction avec les étudiants.

Je suis également fortement engagée dans la définition des UEs en lien avec la programmation, les mathématiques discrète et l'apprentissage automatique. En Iran, j'ai créé du matériel pédagogique pour

les cours de programmation, de mathématiques discrètes, de conception Web et d'ingénierie logicielle. Ces activités s'accompagnaient de la prise de contacts avec les enseignants et de la planification des cours hebdomadaires pour le département d'ingénierie informatique pour chaque semestre. À cette période, j'avais reçu les prix en tant que meilleure enseignante pour certains semestres. Le tableau 1 récapitule mes expériences en enseignement suivant les deux axes de la matière enseignée et du public visé depuis 2010.

	Licence (L1,L2,L3)	Master M1	Master M2	Années
Apprentissage automatique (Master Progis)			35h (20h TD, 15h TP)	2024-25
Fondements des bases de données (M1-MOSIG)		41.25h (TD)		2022-23
Introduction aux bases de données (M1-AM)		36.63h (20.63h CM,16h TP)		2022-23
Informatique pour les sciences de la vie	33h (chaque année) (18h TD,15h TP) - L1			2020-23
Algorithmique	36h (chaque année) CM/TD- L2			2010-19
Programmation (C++)	48h (chaque année) (TD,TP)- L1			2010-19
Mathématiques discrètes	24h (chaque année) TD- L1			2010-19
Ingénierie logicielle	24h (chaque année) CM- L3			2010-19
Conception de sites Web	24h (chaque année) CM,TP- L3			2010-19
Total	819h - L1 ; 360h - L2 ; 480h - L3	77h (5h vacataire)	35h	1771h

TABLE 1 – Tableau synthétique des enseignements. Les sigles CM, TD et TP signifient la nature des enseignements, respectivement : Cours Magistral, Travaux Dirigés et Travaux Pratiques.

Recherche

Thèmes de recherche (mots clés) : **Intelligence artificielle :** apprentissage automatique, traitement du langage naturel, détection d'émotions dans les données textuelles, résumé de textes ; **Ingénierie logicielle :** technologies de développement logiciel, gestion de base de données, web sémantique et ontologie ; **Tests logiciels :** analyse des logs , tâches d'exploration des logs (y compris minimisation des logs, prédiction des échecs, détection des causes profondes des événements), détection des anomalies et des valeurs aberrantes ; **Bioinformatique :** algorithmes bioinformatiques avancés, bases de données biologiques.

Activité de Recherche

J'ai démarré mes travaux de recherche par l'analyse de données et l'examen de bases de données bioinformatiques. Par la suite, après m'être familiarisé avec les journaux logiciels, j'ai commencé à réaliser des recherches sur les tâches d'exploration des journaux logiciels (logs), en particulier la minimisation des logs, la prédiction des échecs et la détection des causes profondes. Je mène actuellement des recherches sur les grands modèles de langage et les agents conversationnels. Je présente maintenant ces trois axes de façon plus détaillée.

— Bioinformatique

En 2017, j'ai entamé mes études doctorales en bioinformatique à l'Université de Téhéran, tout en exerçant comme ingénieure de formation. J'ai suivi environ 35 unités de formation, portant notamment sur les bases de données biologiques, l'exploration de données dans les systèmes de

santé, les algorithmes bioinformatiques avancés, ainsi que les principes de biochimie et de biologie moléculaire, afin de me familiariser avec les concepts de ce domaine. Durant cette période, j'ai participé à divers projets, en particulier à l'analyse de la base de données TCGA et au projet Microbiome. Malgré mon intérêt pour ce domaine, j'ai été contrainte d'abandonner mes études de doctorat en bioinformatique après notre déménagement en France.

- En mars 2019, J'ai travaillé comme ingénieure de recherche au sein du groupe TIMC-IMAG du laboratoire TIMC de l'université de Grenoble sous la direction de Magali Richard (magali.richard@univ-grenoble-alpes.fr). Durant 6 mois, j'ai étudié les données de méthylation de l'ADN en utilisant le langage de programmation R et j'ai travaillé sur certaines approches existantes. J'ai présenté un poster lié à mon travail lors de la première édition de la "Global Challenges Science Week" en juin 2019 à Grenoble.

— Tests de logiciels et apprentissage automatique

En septembre 2019, j'ai commencé mon doctorat dans le domaine des tests logiciels à l'université de Grenoble Alpes. Ma thèse s'est articulée autour du projet Philae (N° ANR-18-CE25-0013). L'objectif de Philae consistait à créer et à maintenir des tests de régression automatisés en réalisant une analyse des traces d'exécution avec les techniques d'apprentissage automatique, et en intégrant cette analyse avec l'inférence de modèle et la génération de tests automatisés.

Plus précisément, je me suis intéressée à l'étude de quelques tâches spécifiques telles que la minimisation des fichiers journaux et des suites de tests [5, 6], la prédiction des échecs [4], la détection des causes profondes des événements et le regroupement des comportements des utilisateurs [3]. Durant ma thèse, j'avais des réunions mensuelles avec les partenaires universitaires de Philae, notamment l'université de Bourgogne-Franche-Comté en France et l'université du Queensland en Australie, ainsi qu'avec des entreprises dont Orange et Smartesting. J'ai été aussi amenée à collaborer avec les membres des deux équipes, Vasco et Aptikal, du LIG. Les trois études de cas que j'ai considérées sont décrites ci-dessous.

— Scannette : l'objectif de cette étude était de réduire d'énormes journaux de suivi des utilisateurs à des journaux plus petits avec le même effet de déclenchement de bugs.

— Télécommunications (Orange) : L'objectif de cette étude était de découvrir les anomalies et de fournir des analyses intelligentes des résultats de l'exécution des tests pour aider les ingénieurs de test à se concentrer sur les sections les plus sujettes aux erreurs du système testé.

— Train Ticket Benchmark : l'objectif de cette étude était d'ajouter une défaillance et de simuler l'utilisation du système en exécutant Stress-ng sur un serveur Docker.

La chaîne de traitement des logs développée pour cette recherche a été publiée sous la forme de deux outils open source distincts, accessibles en ligne ² dans la boîte à outils Philae.

— Intelligence artificielle générative

Depuis mai 2024, je suis engagée dans le projet de création d'un agent conversationnel intégrant un analyseur d'émotions. L'objectif est de développer un chatbot avancé utilisant un modèle de langage de grande taille pour générer automatiquement des questions pertinentes et personnalisées en fonction des interactions de l'utilisateur et de ses émotions. Cette analyse s'effectue sur les données audio et textuelles produits par l'utilisateur. Pour ce travail, je suis co-auteure d'un brevet soumis.

Encadrement de recherche

- 2024-2025 : Co-encadrement d'un ingénieur de recherche dans le cadre du développement d'un agent conversationnel durant mon post-doctorat.
- 2022 - Stage de M1 de 4 mois sur le thème de "Apprentissage automatique pour la détection d'anomalies logicielles", Grenoble INP, ENSIMAG.
- 2022 - Stage de L3 de 2 mois sur le thème de "Analyse de logs pour le test de systèmes logiciels", Grenoble INP, ENSIMAG.

Publications

Le tableau 2 présente une synthèse de mes publications, qui proviennent principalement de ma thèse de doctorat.

De plus, je suis activement impliquée dans le développement d'un agent conversationnel pour mon post-

2. <https://github.com/PHILAE-PROJECT>

Type	Nombre
Ouvrage	1
Conférences internationales	2
Ateliers internationaux	2
Conférences nationales	2

TABLE 2 – Résumé de mes publications

doctorat et un brevet a été déposé dont je suis co-auteur.

L'ensemble des publications est décrit ci-dessous et il est disponible à cette adresse :

Lien vers Google Scholar: <https://scholar.google.com/citations?user=wVYtFx0AAAAJ&hl=en> Lien vers ma page personnelle: <https://afshinpour.com/publications.html>

Références

Thèse

- [1] Bahareh Afshinpour, Mining Software Logs with Machine Learning Techniques, Université Grenoble Alpes, 2023, Sep, Theses, <https://theses.hal.science/tel-04233033>, https://theses.hal.science/tel-04233033v3/file/AFSHINPOUR_2023_diffusion.pdf, tel-04233033, v3.

Ouvrage

- [2] Bahareh Afshinpour, Python programming language training book for students (en Farsi), *Faragir Hegmataneh*, pp. 1-140. 2015, [Link to book \(Faragir Hegmataneh store\)](#).

Conférences internationales avec comité de lecture

- [3] Bahareh Afshinpour, Massih-Reza Amini, et Roland Groz, Semantic Log Partitioning : Towards Automated Root Cause Analysis, in *2024 IEEE 24th International Conference on Software Quality, Reliability and Security (QRS)*, pp. 639–648, 2024, <https://qrs24.techconf.org/>, doi : [10.1109/QRS62785.2024.00069](https://doi.org/10.1109/QRS62785.2024.00069).
- [4] Bahareh Afshinpour, Roland Groz, et Massih-Reza Amini, Telemetry-Based Software Failure Prediction by Concept-Space Model Creation, in *2022 IEEE 22nd International Conference on Software Quality, Reliability and Security (QRS)*, pp. 199–208, 2022, <https://qrs22.techconf.org/>, doi : [10.1109/QRS57517.2022.00030](https://doi.org/10.1109/QRS57517.2022.00030).

Ateliers internationaux avec comité de lecture

- [5] Bahareh Afshinpour, Roland Groz et Massih-Reza Amini. Correlating Test Events With Monitoring Logs For Test Log Reduction And Anomaly Prediction. In *33rd IEEE International Symposium on Software Reliability Engineering Workshops (ISSRE)*, pp. 274-280, 2022, <https://issre2022.github.io/index.html>.
- [6] Bahareh Afshinpour, Roland Groz, Massih-Reza Amini, Yves Ledru et Catherine Oriat. Reducing Regression Test Suites using the Word2Vec Natural Language Processing Tool. In *1st Workshop on Natural Language Processing Advancements for Software Engineering (NLPaSE)*, 2020, <https://sites.google.com/view/nlpase2020/about-nlpase>.

Conférences nationales avec comité de lecture

- [7] Bahareh Afshinpour, Roland Groz et Massih-Reza Amini. Apprentissage collaboratif pour la réduction des log de test et la prédiction des anomalies sur les log de monitoring, In *24^{eme} Conférence Francophone sur l'Apprentissage Automatique (CAp)*, 2022, <https://caprfiap2022.sciencesconf.org/>.
- [8] Bahareh Afshinpour, Hassan Khotanlou, Ahmad Farahi. Un nouvel algorithme de classement basé sur l'ontologie pour les moteurs de recherche Web sémantique (en Farsi), In *3^{eme} conférence iranienne sur l'ingénierie informatique et les technologies de l'information*, 2011, <https://civilica.com/doc/106102/certificate/print/>.