

Hà Nội, ngày 23 tháng 5 năm 2024

THÔNG BÁO SỐ 1

Về việc tổ chức các buổi bồi dưỡng trực tuyến “AI Programming for Beginners” phổ cập kiến thức lập trình Trí tuệ nhân tạo / Học máy cho học sinh các trường phổ thông trung học chuyên trên toàn quốc

Viện Tin học Nhân dân là một tổ chức KH-CN, trực thuộc Hội Tin học Việt Nam (VAIP). Về Đào tạo và Phổ biến kiến thức, Viện có sứ mệnh đào tạo giải pháp, công nghệ, kỹ năng chuyên sâu, phổ cập kiến thức về CNTT-TT cho toàn xã hội.

Hiện nay nhu cầu về phát triển nguồn nhân lực trình độ quốc tế về Trí tuệ nhân tạo (AI) rất cấp thiết nhằm phục vụ Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng AI đến năm 2030. Viện Tin học nhân dân tiến hành tổ chức các buổi bồi dưỡng trực tuyến “AI Programming for Beginners” nhằm phổ cập kiến thức lập trình Trí tuệ nhân tạo / Học máy dành cho học sinh các trường phổ thông trung học chuyên trên toàn quốc, đồng thời qua đây hỗ trợ các trường thực hiện chương trình đổi mới giáo dục phổ thông chuyên 2019. Thông tin chi tiết như sau.

I. Thời gian và hình thức bồi dưỡng

- Khoá bồi dưỡng diễn ra trong 5 tuần từ 27/5 đến 28/6/2024.
- Mỗi tuần gồm ba buổi vào các tối thứ 2, thứ 4 và thứ 6.
- Mỗi buổi gồm 3 giờ học từ 20h – 23h.
- Khoá bồi dưỡng được tổ chức trực tuyến trên nền tảng Zoom.

II. Chuyên gia

Các chuyên gia nhiều kinh nghiệm của Viện Tin học Nhân dân tham gia xây dựng chương trình, giảng dạy và trợ giảng đến từ các trường đại học lớn trong cả nước:

- PGS.TS Đỗ Phan Thuận, Giảng viên cao cấp Trường Đại học Tôn Đức Thắng
- PGS.TS Nguyễn Phi Lê, Giảng viên cao cấp Đại học Bách khoa Hà Nội.
- TS. Nghiêm Nguyễn Việt Dũng, Giảng viên Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.
- KS. Nguyễn Đắc Thái, học viên thạc sĩ Khoa học dữ liệu, VinUni.
- Trịnh Văn Chiến, kỹ sư AI tại Công ty CMC ATI.
- Nguyễn Tuấn Hiệp, Đại học Bách khoa Hà Nội.
- Nguyễn Thành Trung, Đại học Bách khoa Hà Nội.
- và các chuyên gia là cố vấn, khách mời khác.

III. Nội dung các buổi bồi dưỡng

Theo nội dung của buổi học, các giảng viên sẽ giảng dạy các phần lý thuyết và hướng dẫn các bài thực hành đan xen. Dưới đây là kế hoạch tổng quan của các buổi bồi dưỡng, nội dung chi tiết sẽ được giới thiệu trong buổi học đầu tiên.

CHƯƠNG I. NGÔN NGỮ LẬP TRÌNH VÀ FRAMEWORK (4 BUỔI)

Mục đích: Cung cấp cho người học những kiến thức cơ bản về:

- Ngôn ngữ lập trình Python
- Các Framework thường dùng trong lập trình các thuật toán AI
- Các thư viện thường dùng trong lập trình các thuật toán AI

Buổi 1.1: Giới thiệu ngôn ngữ lập trình Python

Buổi 1.2: Giới thiệu các Framework thường dùng trong lập trình các thuật toán AI

Buổi 1.3: Giới thiệu các thư viện thường dùng trong lập trình các thuật toán AI

Buổi 1.4: Một số các bài tập ứng dụng

CHƯƠNG II. MÔI TRƯỜNG LÀM VIỆC (1 BUỔI)

Mục đích: Trang bị các kiến thức cơ bản về:

- Google Colab và các môi trường thường dùng để code các mô hình AI
- Các công cụ thường dùng để trực quan dữ liệu, theo dõi đánh giá quá trình huấn luyện mô hình: wandb, ...

Bài 1. Google Colab và ứng dụng trong phát triển các mô hình AI

Bài 2. Các công cụ trực quan hóa để giám sát quá trình huấn luyện mô hình

CHƯƠNG III. HỌC MÁY CƠ BẢN (3 BUỔI)

Mục đích: Giới thiệu khái niệm AI, học máy. Trang bị các kiến thức về xử lý dữ liệu. Cách sử dụng các mô hình học máy cơ bản.

- Các khái niệm AI, ML, tác vụ, mô hình, dữ liệu, huấn luyện, kiểm thử
- Các dạng tác vụ, các kiểu thuật toán huấn luyện mô hình
- Xử lý dữ liệu (phân tích phân phối dữ liệu, loại nhiễu, loại outlier, điền dữ liệu thiếu, ...)
- Các mô hình học máy cơ bản: hồi quy, KNN, SVM, Decision Tree, Random Forest, XGBoost, ANN

Buổi 3.1: Nhập môn AI/ML và xử lý dữ liệu

Thực hành: Phân tích và làm sạch bộ dữ liệu nhỏ

Buổi 3.2: Các thuật toán học máy cơ bản

Thực hành: Xây dựng 1 số mô hình hồi quy và phân loại

Buổi 3.3: Các thuật toán học máy nâng cao

Thực hành: Áp dụng kỹ thuật Grid Search để tối ưu hóa một mô hình cụ thể

CHƯƠNG IV. THỊ GIÁC MÁY TÍNH (3 buổi)

Mục đích: Trang bị các kiến thức cơ bản trong lĩnh vực thị giác máy tính. Giới thiệu một số mô hình thường dụng trong lĩnh vực thị giác máy tính.

- Các loại bài toán trong lĩnh vực thị giác máy tính (object detection, segmentation) và ứng dụng thực tế.
- Mạng CNN: cơ sở lý thuyết, cách dùng, cách xử lý dữ liệu, ...
- Một số model họ CNN thông dụng: cách sử dụng, các hyperparameters thường finetune.

Buổi 4.1: Giới thiệu về thị giác máy tính

Thực hành: Ứng dụng các thuật toán cổ điển giải quyết bài toán phân loại ảnh chữ số

Buổi 4.2: Mạng CNN - cơ sở lý thuyết và cách dùng

Thực hành: Sử dụng mạng CNN giải quyết bài toán phân loại ảnh.

Buổi 4.3: Một số mô hình CNN thông dụng

Thực hành: Ứng dụng các mô hình CNN giải quyết bài toán khoanh vùng vật thể. Thực hành các kỹ thuật hiệu chỉnh siêu tham số.

CHƯƠNG V. XỬ LÝ NGÔN NGỮ TỰ NHIÊN (4 BUỔI)

Mục đích: Trang bị các kiến thức cơ bản trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP). Giới thiệu về lý thuyết và các mô hình xác suất truyền thống. Nhập môn về mô hình ngôn ngữ lớn, cách sử dụng các mô hình ngôn ngữ lớn.

Buổi 5.1: Giới thiệu về xử lý ngôn ngữ tự nhiên

Thực hành: Làm quen với công cụ NLTK; Làm quen với các bộ dữ liệu; Định dạng đầu vào và đầu ra của từng bài toán; Chạy thử với các mô hình đơn giản.

Buổi 5.2: Giới thiệu về các mô hình xác suất trong NLP

Thực hành: Tính toán về xác suất từ loại theo công thức Markov; Huấn luyện mô hình Markov với tập dữ liệu từ nltk; Xử lý văn bản với tf-idf và ứng dụng vào bài toán tóm tắt văn bản; Thực hành xây dựng n-gram; Thực hành sử dụng không gian vector của từ vựng để so sánh câu.

Buổi 5.3: Giới thiệu về các thư viện lưu trữ mô hình NLP

Thực hành: Thực hành tạo ra pipeline cho tác vụ hỏi đáp với ngôn ngữ tiếng Anh; Thực hành tạo ra pipeline cho tác vụ nhận dạng thực thể có tên với ngôn ngữ tiếng Việt; Cài đặt mô hình cho tác vụ tìm kiếm văn bản (ranking); Tinh chỉnh mô hình với một bộ dữ liệu cho trước.

Buổi 5.4: Giới thiệu về các mô hình ngôn ngữ lớn

Thực hành: Một số mô hình LLM thông dụng; Viết hàm so sánh văn bản dựa trên một mô hình LLM cho trước; Điều chỉnh tập từ vựng và quá trình Tokenization; Học ít mẫu và tinh chỉnh mô hình LLM sẵn có.

IV. Phí tham gia:

- Học viên tự do: $300.000 \text{ VNĐ} / \text{buổi} \times 15 \text{ buổi}$.
- Học sinh đăng ký theo trường hoặc theo giáo viên hướng dẫn môn Tin học: $200.000 \text{ VNĐ} / \text{buổi} \times 15 \text{ buổi}$.

- Học sinh đăng ký tham gia kì thi tuyển chọn đội tuyển Olympic Trí tuệ nhân tạo vòng quốc gia tham gia kì thi Olympic Trí tuệ nhân tạo quốc tế IOAI lần thứ nhất tại Bungary 9-15/8/2024: 150.000 VNĐ / buổi \times 15 buổi. Link đăng kí vòng sơ loại quốc gia (thi trong vòng 1 tháng từ 19/5 đến 20/6) tại website <https://ioai.vn>.
- Thông tin chuyển khoản phí tham gia:
 - Tên tài khoản: Hội Tin học Việt Nam
 - Số tài khoản: 003 70406 0000 979 tại Ngân hàng TMCP Quốc tế Việt Nam (Phòng GD: 64-68 Lý Thường Kiệt, Hoàn Kiếm, Hà Nội, Mã Citad: 01314007)
- Nội dung: Trường/Học sinh xyz nộp lệ phí tham dự khoá bồi dưỡng AI

V. Chứng chỉ

- Các học viên sau khi hoàn thành khoá học sẽ được cấp chứng chỉ hoàn thành khoá học do Viện Tin học Nhân dân - Hội tin học Việt Nam chứng nhận.

VI. Đăng ký

- Thông tin, đăng ký và hướng dẫn đăng ký theo đường link sau :

<https://forms.gle/yQ5kSjTU7QUM9bzq8>

Các cá nhân và Trường hoàn thành chuyển khoản phí tham gia xin vui lòng gửi xác nhận chuyển khoản qua email đến: **OlpVietnam@vaip.vn**, hoặc **office@vaip.vn** trước 27/05/2024.

Thông tin về Khoá bồi dưỡng liên hệ với Viện Tin học Nhân dân - VAIP

Ban Đào tạo - PBKT, tầng 6 số 14 Trần Hưng Đạo, Hà Nội.

Tel: +84 024.38211725; Email: OlpVietnam@vaip.vn; office@vaip.vn

Trân trọng./.

Nơi gửi:

- Các trường chuyên THPT
- Lưu

VIỆN TRƯỞNG



Hoàng Quốc Lập