ĐỀ CƯƠNG HỌC PHẦN

HỌC MÁY TRONG THỦY VĂN VÀ TÀI NGUYÊN NƯỚC

Machine learning in Hydrology and Water resources

1. **Mã học phần**: HMO5183
2. **Số tín chỉ**: 3 (Phân bổ giờ học tập: Lý thuyết/Thực hành/Tự học): 30/30/90
3. **Học phần tiên quyết**: Không
4. **Ngôn ngữ giảng dạy**: Tiếng Việt
5. **Giảng viên**:

STT	Họ và tên	Học hàm, học vị	Đơn vị công tác	Email
1.	Nguyễn Tiền Giang	PGS. TS	Bộ môn Thủy văn và Tài nguyên nước, Khoa Khí tượng Thủy văn và Hải dương học	giangnt@vnu.edu.vn
2.	Trần Anh Phương	Tiến sĩ	Viện Khoa học Tài nguyên nước	phuongtran.monre@gmail.com
3.	Nguyễn Quang Hưng	Tiến sĩ	Bộ môn Thủy văn và Tài nguyên nước, Khoa Khí tượng Thủy văn và Hải dương học	hungnq@hus.edu.vn

6. **Mục tiêu của học phần**:

6.1. Mục tiêu chung: Trang bị cho học viên kiến thức cơ bản cũng như nâng cao về máy học và ứng dụng của máy học để giải quyết các bài toán thường gặp trong lĩnh vực Thủy văn, Tài nguyên nước.

6.2. Mục tiêu cụ thể

a) Kiến thức: Kết thúc học phần, học viên có kiến thức cơ bản về máy học và ứng dụng của máy học trong lĩnh vực thuỷ văn tài nguyên nước; hiểu được nguyên lý và cách sử dụng các mô hình và thuật toán máy học phổ biến áp dụng trong bài toán thực tiễn; nhận diện được các vấn đề có thể ứng dụng máy học cũng như các thách thức và khó khăn khi sử dụng phương pháp này trong lĩnh vực thuỷ văn tài nguyên nước. Ngoài ra, học viên sẽ tiếp thu các kiến thức tổng quát về các lĩnh vực khác bổ trợ cho máy học như xác suất thống kê và lập trình ngôn ngữ bậc cao.

b) Kỹ năng: Học viên được rèn luyện kỹ năng tự học, tự nghiên cứu sử dụng phương pháp máy học giải quyết các bài toán cụ thể thông qua việc tìm kiếm, thu thập, tổng hợp, phân tích các tài liệu liên quan; có kỹ năng viết và trình bày một vấn đề khoa học cụ thể.

c) Mức tự chủ và trách nhiệm: Rèn luyện đức tính trung thực trong nghiên

cứu, có trách nhiệm với các kết quả, sản phẩm của công việc, sẵn sàng chia sẻ kiến thức, tôn trọng bản quyền; có khả năng thích ứng trong các môi trường làm việc độc lập cũng như làm việc nhóm.

7. **Tóm tắt nội dung học phần:**

Học phần cũng cấp các ứng dụng máy học vào giải quyết các bài toán thực tiễn trong lĩnh vực thuỷ văn và tài nguyên nước. Học phần chú trọng và phát triển kỹ năng thực hành và kiến thức nền tảng để thực hiện nghiên cứu bằng phương pháp máy học, cung cấp cho học viên kinh nghiệm thực tế sử dụng các công cụ máy học mã nguồn mở hiện đại. Học viên sẽ được thực hành các bước cơ bản trong phát triển và ứng dụng máy học, bao gồm nhận diện vấn đề, phân tích dữ liệu, tìm giải pháp tối ưu, sử dụng các hàm thống kê và thuật toán cao cấp xác định mối quan hệ giữa các đại lượng nghiên cứu và diễn giải kết quả. Học viên sẽ học sử dụng ngôn ngữ lập trình Python và thư viện Scikit-learn để xây dựng thuật toán máy học và áp dụng vào thực tiễn lĩnh vực thuỷ văn tài nguyên nước.

8. **Chuẩn đầu ra của học phần**:

8.1. Kiến thức (*Course Knowledge* – CK)

- CK1: Hiểu được các khái niệm cơ bản và nguyên lý áp dụng phương pháp máy học trong thuỷ văn tài nguyên nước (mức 2)

- CK2: Phân biệt sự khác nhau và khả năng ứng dụng của các mô hình máy học áp dụng đối với các vấn đề cụ thể trong thuỷ văn học. (mức 4)

- CK3: Lựa chọn và áp dụng các mô hình hoặc thuật toán máy học phù hợp để giải quyết các bài toán thuỷ văn tài nguyên nước theo cách tiếp cận mới (mức 4)

- CK4: Đánh giá và phân tích được các kết quả từ các mô hình máy học. (mức 5)

8.2. Kỹ năng (*Course Skill* – CS)

- CS1: Học viên lựa chọn được thông tin từ các nguồn khác nhau liên quan đến một chủ đề cần thảo luận, nghiên cứu dựa trên tư duy phản biện (mức 3).

- CS2: Học viên phát hiện, nhận diện được các vấn đề phức tạp có liên quan đến hệ thống thủy văn theo tư duy hệ thống (mức 3).

- CS3: Học viên Kết hợp nhuần nhuyễn các kỹ năng giao tiếp, kỹ năng truyền tải thông tin (mức 4).

8.3. Mức tự chủ và trách nhiệm (*Course Responsibility* – CR)

- CR1: Thích ứng với các yêu cầu làm việc độc lập và làm việc theo nhóm trong các điều kiện khác nhau. (mức 4)

- CR2: Trung thực trong nghiên cứu, có trách nhiệm với các kết quả, sản phẩm của công việc, sẵn sàng chia sẻ kiến thức, tôn trọng bản quyền. (mức 4)

9. **Ma trận liên kết chuẩn đầu ra của học phần với chuẩn đầu ra của CTĐT**

(Mức độ đóng góp đánh giá theo thang Bloom của học phần vào các chuẩn đầu ra của CTĐT theo Kiến thức (Program Knowledge-PK), Kỹ năng (Program Skill-PS), Mức tự chủ và trách nhiệm (Program Responsibility-PR)

CĐR CTĐT	**PK3**	**PK4**	**PK5**	**PS1**	**PS2**	**PS3**	**PS4**	**PS5**	**PS6**	**PR1**	**PR2**
Mức đóng góp của học phần	4	4			4	4		4			

10. **Kiểm tra, đánh giá kết quả học tập**

STT	**Tiêu chí đánh giá**	**Trọng số**	**Mô tả**	**Chuẩn đầu ra cần đo được**
1	Thường xuyên	20%	*Đánh giá điểm chuyên cần*	CK1, CS1, CR1, CR2
2	Giữa kỳ	20%	*Thi viết hoặc viết báo cáo và trình bày dự án*	CK2, CK3, CS1, CS2 , CR1, CR2
3	Cuối kỳ	60 %	*Thi viết, thi vấn đáp hoặc viết báo cáo và trình bày dự án*	CK3, CK4, CS1, CS3, CR2
	Tổng	100%		

11. **Học liệu**

11.1 Học liệu bắt buộc (tác giả, tên giáo trình, nhà xuất bản, năm xuất bản):

- Nguyễn Tiền Giang. Tài liệu trình chiếu học phần "Máy học trong Thủy văn và Tài nguyên nước", ĐHKHTN (cập nhật hàng năm).
- Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2019.

11.2 Học liệu tham khảo (tác giả, tên giáo trình, nhà xuất bản, năm xuất bản):

- Nelli F. Python data analytics: with pandas, numpy, and matplotlib, 2nd edn. Apress, New York, US, 2018
- William W. Hsieh (2009), Machine Learning Methods in the Environmental Sciences Neural Networks and Kernels, Cambridge University Press.

12. **Nội dung chi tiết học phần và liên kết nội dung với chuẩn đầu ra**

(*lưu ý: Nội dung chi tiết thể hiện chuẩn đầu ra của học phần, có thể chi tiết thông tin về phân bổ giờ học tập, lịch trình giảng dạy*)

Chương/ mục	Tiêu đề/Nội dung	Chuẩn đầu ra của học phần
Chương 1. Giới thiệu chung		
1.1	Máy học là gì và sao lại sử dụng máy học?	
1.2	Một số ví dụ	
1.3	Các cách phân loại	
1.4	Các thách thức chính trong máy học	
1.5	Kiểm thử và kiểm định	
1.6	Thực hành 1 – Lập trình ngôn ngữ Python	
Chương 2. Thiết lập một dự án máy học		
2.1	Làm việc với dữ liệu thực tế	
2.2	Nhìn tổng thể	
2.3	Thu nhận dữ liệu	
2.4	Khám phá và trực quan hóa dữ liệu	
2.5	Chuẩn bị dữ liệu	
2.6	Lựa chọn và huấn luyện mô hình	
2.7	Cải thiện các mô hình và tối ưu hóa các siêu tham số	
2.8	Thực hành 2 – Thiết lập một dự án máy học cho một bài toán thủy văn	
Chương 3. Huấn luyện các mô hình máy học		
3.1	Hồi quy tuyến tinh	
3.2	Gradient Descent	
3.3	Hồi quy hàm đa thức	
3.4	Các đường cong huấn luyện	
3.5	Các mô hình tuyến tính điều chuẩn (regularized)	
3.6	Hồi quy logistic	
3.7	Thực hành 3 – Xây dựng mô hình quan hệ mực nước – lưu lượng	
Chương 4. Các mô hình máy véc tơ hỗ trợ (SVMs)		
4.1	Mô hình SVM phân loại tuyến tính	
4.2	Mô hình SVM phân loại phi tuyến	
4.3	Mô hình SVM hồi quy tuyến tính	
4.4	Một số vấn đề chuyên sâu	
4.5	Thực hành 4 – Thí nghiệm với các hàm kernel trong bài toán mô phỏng mực nước, độ mặn	

Chương 5. Các cây quyết định, Học tổ hợp và Rừng ngẫu nhiên		
5.1	Huấn luyện, trực quan hóa, dự báo với Cây quyết định	
5.2	Giải thuật huấn luyện CART sử dụng Scikit-Learn	
5.3	Điều chuẩn các cây và sử dụng cho các bài toán hồi quy	
5.4	Các hạn chế của mô hình cây quyết định	
5.5	Các mô hình tổ hợp (voting classifiers)	
5.6	Bagging and Pasting trong Scikit-Learn	
5.7	Các mô hình rừng ngẫu nhiên	
5.8	Boosting và Stacking	
5.9	Thực hành 5 – Dự báo đóng mở cửa sông	
Chương 6. Mô hình mạng nơ-ron truyền thẳng (feed-forward)		
2.1	Mô hình McCulloch và Pitts	
2.2	Mô hình perceptron	
2.3	Mô hình perceptron nhiều lớp (MLP)	
2.4	Giải thuật lân truyền ngược	
2.5	Các nơ-ron ẩn	
2.6	Các hàm radial basis	
2.7	Thực hành 6 – Mô hình máy học mưa dòng chảy	

Trưởng Khoa	**Trưởng Bộ môn**	**Giảng viên**
Công Thanh	**Nguyễn Tiền Giang**	**Nguyễn Tiền Giang**