**INTRODUCTION TO DATA SCIENCE**

Giảng viên : TS. Bùi ThÔng Hùng

Bộ môn Khoa học dữ liệu, Khoa Công nghệ thông tin

Đại học Công nghiệp thành phố Hồ Chí Minh

Email : [buithÔnghung@iuh.edu.vn](mailto:buithanhhung@iuh.edu.vn)

Website: [https://sites.google.com/site/hungthÔngbui1980/](https://sites.google.com/site/hungthanhbui1980/)

Sinh viên thực hiện : Trịnh Dương Hoan

MSSV : 22684251

Lớp : DHKHDL18A

Bài Làm

**Bài 1:**

**Đánh giá bản thân thông qua SWOT**

|  |  |
| --- | --- |
| Strengths ( Điểm mạnh ) | Weaknesses ( Điểm yếu ) |
| * Kĩ năng giao tiếp tốt * Siêng năng, có thể ngồi học và nghiên cứu trong thời gian lớn trong ngày * Khả năng tiếp thu kiến thức mới khá nhÔng * Tư duy khá tốt | * Tiếng Ông còn hạn chế * Khả năng teamwork chưa được tốt, những vấn đề đã từng làm nhóm chưa thực sự tối ưu * Chưa hình thành thói quen lập kế hoạch chi tiết cho từng ngày, tuần, tháng , năm ,.. |
| Opportunities ( Cơ hội ) | Threats ( Thách thức) |
| * Đang được học tại một trong những trường top đầu về cntt ở miền Nam * Được tham gia các câu lạc bộ về học thuật như Programming Lab và AI Lab * Được tiếp xúc với các Ông chị có kinh nghiệm đi trước , tiếp xúc với những bài báo, đồ án mà các Ông chị đã làm qua để học hỏi kinh nghiệm | * Kiến thức chuyên môn còn hạn chế, chưa có nhiều cơ hội để tiếp cận với những vấn đề thực tế * Công nghệ thông tin đang là một trong những ngành hot ở hiện tại, đặc biệt là lĩnh vực data, vì vậy việc cạnh trÔng trong công việc là điều không thể tránh khỏi, vì thế phải rèn luyện kiến thức chuyên ngành và khả năng ngoại ngữ tốt để nổi bậc hơn các ứng viên còn lại |

**Kế hoạch cho 5 năm sắp tới**

* Mục tiêu : Ra trường đúng hạn với tấm bằng loại giỏi , tích lũy đầy đủ kiến thức chuyên môn lẫn kĩ năng, trở thành một Data Engineer xuất sắc.
* Kế hoạch cụ thể :
* 2024
* Tích lũy kiến thức ở trường ở mức tốt
* Học tiếng Ông ( >= 1500 từ )
* Giải nhì OLP Tin học chuyên Tin
* Giải Ba ICPC regional
* Có 1 bài báo được chấp nhận ở hội nghị khoa học trẻ ( YSC)
* Tham gia các cuộc thi về Data/AI và có giải
* Có học bổng khuyến khích học tập
* 2025
* Tích lũy kiến thức ở trường ở mức tốt
* Học tiếng Ông , cố gắng đạt được IELTS 6.5
* Tham gia YSC, có bài báo được đăng ở các hội nghị lớn hơn như Eureka, ..
* Tham gia các cuộc thi về Data/AI và có giải
* Có học bổng
* Đi thực tập
* 2026
* Tốt nghiệp với tấm bằng loại giỏi
* Có việc làm ổn định trong ngành Data Engineer
* Giao tiếp tiếng Ông như người bản xứ
* 2027
* Thăng tiến trong công việc
* Cải thiện khả năng tiếng Ông (Mục tiêu là 7.5 IELTS)
* Chăm sóc gia đình nhiều hơn
* 2028
* Có người yêu ( đây là một trong những mục tiêu trước giờ em chưa làm được :vv)
* Thăng tiến trong công việc
* Nghiên cứu thêm về lĩnh vực AI

**Bài 2**

**Bài Làm**

Bài báo của GS. Hồ Tú Bảo về Cách mạng Công nghiệp lần thứ tư và mở ra một góc nhìn sâu sắc về vai trò quyết định của dữ liệu trong sự phát triển của các quốc gia. GS. Bảo trình bày vấn đề mối đe dọa là “**Chỉ một số ít quốc gia thắng cuộc sẽ lấy tất**” và ông đã đặt ra câu hỏi quan trọng về vị thế của Việt Nam trong cuộc đua khóc liệt này.

Trong bài báo, GS. Hồ Tú Bảo nhấn mạnh yếu tố cơ bản của Cách mạng công nghiệp lần thứ 4 là sản xuất thông minh dựa trên công nghệ số và đánh giá rằng quốc gia nào kiểm soát được nguồn dữ liệu lớn và phức tạp này sẽ thắng cuộc. Ông chia sẻ về các chương trình chiến lược quốc gia của các nước tiên tiến như Mỹ, Đức, Pháp, Hàn Quốc, Trung Quốc và Singapore, tập trung vào số hóa, kết nối và phân tích dữ liệu lớn, với ưu tiên cho Khoa học Dữ liệu (KHDL).

GS. Bảo mô tả đặc điểm chính của KHDL và nhấn mạnh vai trò của nó trong quản lý và phân tích dữ liệu để tạo ra hiểu biết và tri thức hành động. Ông thể hiện rằng sự thăng tiến trong công nghệ phân tích dữ liệu đã tạo ra những đột phá lớn và đề cập đến những ứng dụng thực tế như dự đoán nhu cầu thị trường, quản lý ngập lụt, và chăm sóc sức khỏe.

GS. Hồ Tú Bảo đề cập đến những bước mà Hàn Quốc và Trung Quốc đã thực hiện để đẩy mạnh KHDL, bằng cách thành lập Viện Dữ liệu Lớn và các chương trình quốc gia. Ông nhấn mạnh sự quan trọng của việc xây dựng nguồn tài nguyên dữ liệu và thách thức mà Việt Nam phải đối mặt trong cuộc đua này.

Cuối cùng, GS. Hồ Tú Bảo đặt ra câu hỏi quan trọng về khả năng của Việt Nam trong Cách mạng công nghiệp lần thứ 4. Ông nhận thức được những tiềm năng của Việt Nam trong lĩnh vực giáo dục toán học và sức mạnh của lực lượng làm về công nghệ thông tin. Tuy nhiên, Ông cũng nhấn mạnh rằng để "thắng chính mình," Việt Nam cần có chính sách, nỗ lực xây dựng nguồn tài nguyên dữ liệu và sự kết nối lực lượng khoa học và công nghệ trong và ngoài nước.

Trong bối cảnh cuộc cách mạng khoa học và công nghệ đang diễn ra, GS. Hồ Tú Bảo kết luận bài viết bằng câu hỏi tư duy: liệu Việt Nam có thể thay đổi cách nghĩ, cách làm, và tìm ra con đường để phát triển hiệu quả trong cách mạng công nghiệp lần thứ 4 không không? Bài viết kết thúc với một lời nhấn mạnh về sức mạnh của giáo dục và động lực của thế hệ trẻ Việt Nam trong việc định hình tương lai của đất nước.

**Bài 3 :**

Các nhóm dữ liệu chính được nhà nước quản lý :

* Dữ liệu hạ tầng
* Dữ liệu dân số và xã hội
* Dữ liệu kinh tế
* Dữ liệu về an ninh và quản lí sự cố
* Dữ liệu về y tế
* Dữ liệu về kinh phí
* Dữ liệu về giáo dục

Các ứng dụng thông minh hoạt động trên các bộ dữ liệu trên :

1. Hệ thống Giáo dục thông minh : Xây dựng mô hình dữ liệu để đánh giá chất lượng giáo dục và đề xuất cải tiến
2. Hệ thống giao dịch và tài chính : Phân tích dữ liệu tài chính để dự đoán xu hướng kinh tế và đưa ra quyết định thông minh về tài chính đô thị. Ứng dụng công nghệ trực quan hóa dữ liệu để xây dựng bộ trình diễn dữ liệu và kết quả dự báo xu hướng phát triển của các chỉ số kinh tế.
3. Hệ thống dự báo an ninh : Phát hiện các mô hình hành vi bất thường để cảnh báo về an ninh đô thị bằng các tính năng như : nhận diện khuôn mặt, nhận dạng các loại phương tiện, phát hiện đám đông, các sự cố về giao thông, an ninh trật tự, …