

Người Việt thể hiện rõ tố chất làm Khoa học dữ liệu qua dịp Tết Nguyên Đán

Tên đầy đủ của bài viết này là Người Việt thể hiện rõ tố chất làm Data Engineering, Data Analytics, Data Science, Artificial Intelligent Engineering – AI Engineering (tiếng Việt tương ứng là Công nghệ dữ liệu, Phân tích dữ liệu, Khoa học dữ liệu, Công nghệ Trí thông minh nhân tạo hoặc gọi sang hơn là Công nghệ Trí tuệ nhân tạo).

Cần nói sơ qua một chút về các công việc này để các bạn chưa quen có thể hình dung một chút:

Data Engineering là công việc chính yếu của người Data Engineer (Kỹ sư dữ liệu). Cụ thể là công việc liên quan đến tổ chức, lưu trữ đầy đủ chính xác, đảm bảo dữ liệu sẵn sàng để truy xuất bởi người phù hợp với các yêu cầu hợp lý về chi phí, tốc độ, bảo mật, và vài yếu tố các tùy theo đặc thù của tổ chức.

Data Analytics là công việc chính yếu của người làm Data Analyst (Phân tích dữ liệu). Cụ thể là công việc xử lý, đảm bảo dữ liệu thỏa mãn yêu cầu đầu vào cho việc phân tích. Công việc phân tích với mục đích là tổng hợp, báo cáo, trực quan hóa dữ liệu để các bên liên quan sử dụng; hoặc lý giải, tìm câu trả lời cho câu hỏi mà ai đó đặt ra (thường là Data Scientist – xem đoạn tiếp theo). Đôi lúc Data Analyst đóng vai trò là Data Scientist tự đặt ra câu hỏi và đi tìm câu trả lời từ dữ liệu luôn. Câu hỏi về dữ liệu thì rất đa dạng, đơn giản như là thống kê, xác suất thông thường; phức tạp hơn một chút là tìm mối quan hệ của các dữ liệu; cao cấp hơn là phải áp dụng các mô hình để dự báo, tiên lượng.

Data Science: người làm công việc này gọi là Data Scientist. Có thể nói là Data Scientist làm được việc Data Analytics. Nhiệm vụ quan trọng của Data Scientist là đặt ra câu hỏi, tức là nêu ra đề bài. Tùy theo môi trường làm việc hoặc nghiên cứu thì câu hỏi sẽ liên quan đến lĩnh vực, nghề nghiệp, hoặc dự án mà Data Scientist được giao nhiệm vụ. Nói tóm lại có thể xem Data Scientist cao hơn Data Analyst một bậc.

AI Engineering: tương tự như trên bạn hoàn toàn có thể suy đoán người làm việc này (các công nghệ về Trí thông minh nhân tạo) gọi là AI Engineer (kỹ sư AI). Về cơ bản AI Engineering là sự kết hợp công việc của Data Analytics với các dữ liệu có tính đặc thù mà do các thiết bị, máy móc nói chung tạo ra. Thông thường các thiết bị này liên quan đến tự động hóa để giúp năng suất con người làm việc tốt hơn, hoặc là bắt chước con người. Ví dụ các dữ liệu liên quan đến âm thanh, giọng nói của con người (bắt chước cái miệng, cái tai con người); hình ảnh của camera, webcam (bắt chước đôi mắt con người); các dữ liệu thu thập từ các cảm biến để đo nhiệt độ, độ ẩm,...từ môi trường, từ chính cơ thể con người,...Nói chung công việc AI có thể hình dung đơn giản là sự kết hợp của Data Analytics có sự tham gia của phần cứng.

Quay lại tiêu đề của bài viết là liên quan đến dịp Tết: Các bạn tự hỏi là tôi mô tả sơ sơ 4 công việc, tương ứng với bốn nhóm nhân sự về AI ở trên thì có liên quan gì đến

Tết của chúng ta? **Bạn bắt đầu có tố chất của một Data Scientist rồi đấy!** Đọc tiếp phần dưới xem xem có phải không nhé!

Mỗi độ xuân về thì các trò chơi vừa mang tính dân gian, vừa có tính khoa học như Bầu cua tôm cá, Xóc đĩa (hoặc Tài xỉu tùy vùng miền), các thẻ loại bài (2 lá – Xì dách, 3 lá – Cào, 6 lá – Cắt Tê, 13 lá – Tiến lên, 9~10 lá - Phỏm).

Nếu bạn chơi Tiến lên thì việc sắp xếp 13 lá bài sao cho thuận tiện để rút khi đánh, và biết mình còn những nhóm bài nào. Lúc đầu chơi lúc phải xòe bài ra để nhớ hết, và gần về cuối phải xếp lại để dấu bài không cho đối phương biết bạn còn mấy lá. Hành động xếp bài có thứ tự, gọn gàng, cần lấy con bài nào, nhóm bài nào một cách thuận tiện cho thấy bạn có tố chất của **Data Engineer**.

Nếu bạn tham gia chơi Xì dách, đặc biệt là làm cái; hoặc chơi Cắt tê 6 lá; hoặc Tiến lên thì chắc chắn bạn phải vận dụng trí não để đoán đối phương đang cầm những loại bài nào. Thậm chí bạn còn phải phân tích tâm lý đối phương. Tổng hợp rất nhiều yếu tố cho thấy bạn có **nền tảng về Xác suất thống kê, phân tích đa biến** để đưa ra quyết định.

Trong môn Bầu cua tôm cá nếu cho bạn lựa chọn là người chơi hoặc làm cái mà bạn chọn làm cái thì chắc chắn bạn có nền tảng tốt về **Xác suất thống kê**. Mỗi ván chơi mà người chơi đủ đông đặt đủ 6 con thì bạn xóc làm sao mà ra cặp đôi hoặc bộ 3 là xác suất bạn lời là rất lớn. Trong một số trường hợp người chơi có thể cố lén nhìn khe hở của người xóc (vì vui Xuân nên họ cố tình để hở 1 tí cho người chơi nào thích nhìn thì nhìn), khi thoáng thấy mặt trước của cục Bầu cua tôm cá thì bạn có thể đoán được mặt trên. Đây chính là khả năng vận dụng suy luận của Hồi qui tuyến tính đơn giản. Đặc biệt hơn nếu bạn nhìn độ rung tay của nhà cái khi xóc, âm thanh lục cục lục cục (3 cục giấy trong cái hộp nhôm khi lắc) rồi bạn đoán mức độ xoay của các cục bầu cua tôm cá thì cho thấy bạn có nền tảng phân tích âm thanh, cảm nhận được độ rung như là tín hiệu của sensor – đây là minh chứng cho tài năng của **AI Engineer**.

Trong tất cả các môn thì có thể chơi xong thì bạn hân hoan vì đã thắng, đôi lúc thắng lớn. Đặc biệt là lúc thua thì trong đầu bạn sẽ đặt ra rất nhiều câu hỏi đại loại như: tại sao hôm nay mình đen thế nhỉ? Tạo sao ván đó mình đỉnh ninh chắc chắn trúng nên đã xuống tiền mà lại ... Đấy! bạn liên tục đặt câu hỏi trong một niềm cảm xúc dù cho cảm xúc không mấy tốt chứng tỏ bạn có năng lực tiềm ẩn của một **Data Scientist**.

Tóm lại nếu bạn là người mong muốn làm trong lĩnh vực Data Engineering/Data Analytics/Data Scientist/AI Engineering thì hãy tranh thủ nhân dịp Tết này thử phát huy những tố chất này nhé. Biết đâu ăn Tết xong, cơ hội lại rộng mở cho các bạn! Nhớ là chơi nhỏ để thể hiện tố chất thôi nhé, không nên lấy ăn thua làm trọng.