**Nội dung làm slide ngày 26/8**:

* **Phần mở đầu**

Khai phá dữ liệu là một ngành học có tính ứng dụng, áp dụng vào thực tiễn cao, khai thác dữ liệu đã đạt được thành công lớn trong nhiều các ứng dụng. Không thể liệt kê tất cả các ứng dụng mà khai thác dữ liệu đóng vai trò quan trọng.

Để chứng minh tầm quan trọng của các ứng dụng như một khía cạnh chính trong nghiên cứu và phát triển khai thác dữ liệu, chúng tôi thảo luận ngắn gọn về hai ứng dụng rất thành công và phổ biến của khai thác dữ liệu: kinh doanh thông minh và các công cụ tìm kiếm.

* **Phần 1 : Kinh doanh thông minh**

**Business intelligence (BI)** cung cấp lịch sử, hiện tại và các dự báo về hoạt động kinh doanh. Ví dụ bao gồm báo cáo, phân tích trực tuyến xử lý, quản lý hiệu suất kinh doanh, thông minh cạnh tranh, nhập điểm chuẩn và phân tích dự đoán.

\*Kết luận

- Khai thác dữ liệu là cốt lõi của kinh doanh thông minh.

* Kỹ thuật phân loại và dự đoán là cốt lõi của phân tích dự đoán trong kinh doanh thông minh, trong đó có nhiều ứng dụng trong việc phân tích thị trường, nguồn cung cấp và bán hàng.
* Phân cụm đóng vai trò trung tâm trong mối quan hệ với khách hàng quản lý, phân nhóm khách hàng dựa trên những điểm tương đồng của họ. Sử dụng các kỹ thuật khai thác đặc trưng, ​​chúng tôi có thể hiểu rõ hơn về các đặc điểm của từng nhóm khách hàng và phát triển các chương trình thưởng khách hàng tùy chỉnh.
* **Phần 2 : Công cụ tìm kiếm**

Công cụ tìm kiếm Web là một máy chủ máy tính chuyên dụng để tìm kiếm thông tin trên Web. Một số công cụ tìm kiếm cũng tìm kiếm và trả về dữ liệu có sẵn trong cơ sở dữ liệu công cộng hoặc mở trực tiếp.

Công cụ tìm kiếm khác với thư mục web ở chỗ các thư mục web được duy trì bởi người biên tập trong khi các công cụ tìm kiếm hoạt động theo thuật toán hoặc bởi sự kết hợp của đầu vào của con người và thuật toán còn các công cụ tìm kiếm web về cơ bản là các ứng dụng khai thác dữ liệu rất lớn.

Dữ liệu khác nhau, kỹ thuật khai thác được sử dụng trong tất cả các khía cạnh của công cụ tìm kiếm

* Thu thập thông tin (ví dụ: quyết định trang nào nên được thu thập thông tin và tần suất thu thập thông tin)
* Lập chỉ mục (ví dụ: chọn các trang được lập chỉ mục và quyết định mức độ chỉ mục nên được xây dựng)
* Tìm kiếm (ví dụ: quyết định cách xếp hạng các trang, phần quảng cáo nào nên được thêm vào và cách kết quả tìm kiếm có thể được cá nhân hóa hoặc thực hiện "Nhận biết ngữ cảnh").

*Thứ hai*, các công cụ tìm kiếm web thường phải xử lý dữ liệu trực tuyến

Để làm điều này, nó có thể xây dựng bộ phân loại truy vấn chỉ định truy vấn tìm kiếm cho các danh mục được xác định trước dựa trên chủ đề truy vấn

Cho dù một mô hình được xây dựng ngoại tuyến, ứng dụng của mô hình trực tuyến phải đủ nhanh để trả lời các truy vấn của người dùng trong thời gian thực.

Một thách thức khác là duy trì và từng bước cập nhật một mô hình trên các luồng dữ liệu đang phát triển nhanh chóng.

*Thứ ba*, các công cụ tìm kiếm web thường phải đối phó với các truy vấn chỉ được hỏi số lần nhỏ.