## **3.1 Giới thiệu bài toán**

Mặc dù chưa có số liệu thống kê cụ thể, nhưng có thể thấy rằng hầu hết học sinh sau khi hoàn tất chương trình lớp 12 đều tiếp tục lựa chọn cho mình những trường Đại hoc, Cao đẳng, Trung cấp. Thông qua việc chọn trường, chọn ngành nghề sẽ giúp các em có được kiến thức, bằng cấp, kỹ năng làm việc để đáp ứng nhu cầu lao động của xã hội sau khi các em tốt nghiệp. Việc chọn đúng ngành nghề giúp các em định hướng đi phù hợp với khả năng của bản thân mình và tránh khỏi những lựa chọn vội vàng để rồi phải bỏ lỡ rất nhiều cơ hội mà đáng ra nếu chọn, định hướng sớm và đúng các em sẽ thành công. Do rất nhiều yếu tố khách quan, chủ quan khác và đa phần các em ở các tỉnh thành, việc tiếp cận các thông tin về chọn ngành nghề cũng hạn chế và thiếu thông tin để lựa chọn và định hướng cho mình.

Việc chọn đúng ngành học trong bối cảnh kinh tế, xã hội hiện nay nhằm tránh việc có nhiều cá nhân lựa chọn sai nghề sẽ dẫn tới giảm sút chất lượng đào tạo, gây lãng phí cho công tác đào tạo và đào tạo lại. Chất lượng nguồn nhân lực sau đào tạo không đảm bảo dẫn tới năng suất lao động không cao, nảy sinh nhiều xáo trộn cho hoạt động của các tổ chức, doanh nghiệp bởi các hiện tượng như: bỏ nghề, chuyển nghề, … Các doanh nghiệp mất thêm chi phí đào tạo và đào tạo lại cho đội ngũ của mình.

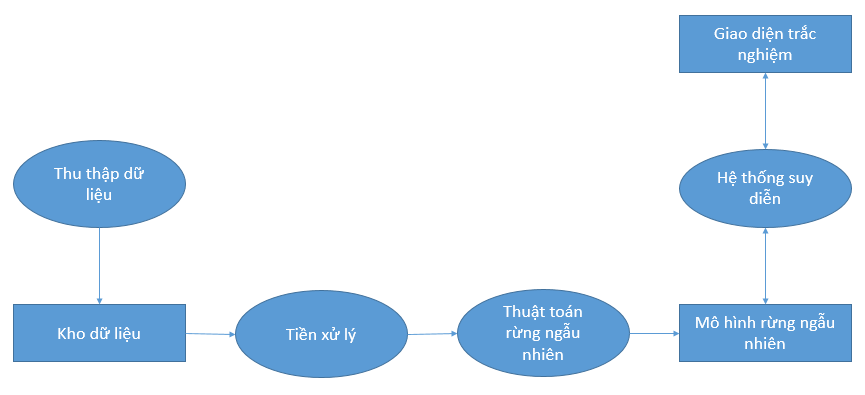
Học tập là một quá trình dài và khó khăn, đòi hỏi sinh viên phải đầu tư nhiều công sức, tiền bạc của cá nhân cũng như gia đình, nếu thất bại trong việc chọn ngành nghề sẽ dẫn đến những tổn thất lớn về cả tinh thần lẫn vật chất cho sinh viên và gia đình họ. Việc định hướng chọn đúng ngành nghề sẽ giúp các em tránh khỏi những vấn đề trên.

Vì vậy, công tác hướng nghiệp chọn ngành nghề cho học sinh trung học phổ thông là điều hết sức cần thiết.

Trước những thực tế đó, tôi xây dựng chương trình TƯ VẤN CHỌN NGHỀ CHO HỌC SINH TRUNG HỌC PHỔ THÔNG trên nền tảng web.

Chương trình được xây dựng nhằm mục đích thông qua các câu hỏi trắc nghiệm định hướng nghề nghiệp của John Holland[3][4] và sử dụng thuật toán học máy để giúp học sinh trung học phổ thông có thể tự định hướng và lựa chọn đúng ngành học phù hợp với năng lực, tính cách và sở thích của bản thân.

## **3.2 Mô hình cấu trúc hệ thống**



Hình 3.1: Mô hình cấu trúc hệ thống tư vấn chọn ngành

Dữ liệu khảo sát thực tế việc chọn ngành học của sinh viên thông qua bảng câu hỏi trắc nghiệm chọn ngành nghề của Tiến sỹ Tâm lý học John Holland[3][4], nhằm phát hiện ra các đặc điểm chung của học sinh, sinh viên trong việc lựa chọn ngành nghề học tập tại trường Đại học Quy Nhơn.

John Holland đã xây dựng một bộ câu hỏi dành cho người muốn tự tìm hiểu mình. Qua nhiều năm phát triển, bộ trắc nghiệm này giúp cho người ta tự phát hiện được các kiểu người trội nhất đang tiềm ẩn trong con người mình để tự định hướng khi lựa chọn nghề. Lý thuyết này dựa trên 8 luận điểm, trong đó 2 luận điểm đầu là: Hầu như ai cũng có thể được xếp vào 1 trong 6 kiểu người, 6 kiểu người đó là Realistic (người thực tế, viết tắt là R), Investigative (người thích nghiên cứu – I), Artistic (người có tính nghệ sĩ – A), Social (người có tính xã hội – S), Enterprising (người dám nghĩ dám làm – E) và Conventional (người công chức – C); có 6 môi trường hoạt động ứng đúng với 6 kiểu người kể trên. Lý thuyết này về sau lấy 6 chữ cái ghép lại thành tên RIASEC[3][4]. Học thuyết của Holland đã lập luận rằng: “Thiên hướng nghề nghiệp chính là sự biểu hiện cá tính của mỗi con người” và nó được phân loại thành 6 nhóm và được diễn tả ở hai phương diện: tính cách con người và môi truờng làm việc. Thuyết Holland không giả định rằng một người chỉ có một trong 6 loại tính cách trên thế giới. Thay vào đó, ông chỉ ra rằng bất kỳ người nào cũng có thể được mô tả bằng việc dung hòa một trong 6 loại tính cách theo thứ tự giảm dần.

Sau quá trình khảo sát chúng ta có một kho dữ liệu gồm bộ kết quả trắc nghiệm và ngành nghề đang học tương ứng. Sau khi có kho dữ liệu, chúng ta tiến hành tiền xử lý dữ liệu, loại bỏ dữ liệu không phù hợp, xác thực thông tin dữ liệu, thay đổi số chiều dữ liệu nhằm tăng độ chính xác khi xây dựng mô hình. Tiếp theo, chúng ta xây dựng thuật toán rừng ngẫu nhiên và sử dụng bộ dữ liệu đã tiền xử lý để huấn luyện và xuất ra mô hình phân loại các ngành nghề dựa theo bộ câu hỏi trắc nghiệm. Từ đó có thể dự đoán được ngành nghề phù hợp khi đưa vào kết quả trắc nghiệm từ người dùng.

Sau khi có được mô hình, chúng ta tạo một ứng dụng web lấy kết quả trắc nghiệm từ người dùng, thông qua hệ thống suy diễn sẽ đưa kết quả trắc nghiệm vào mô hình từ đó lấy ngành nghề phù hợp rồi hiển thị lên trang web cho người dùng.