
Giáo viên:

Vũ Quốc Hoàng – Nguyễn Văn Quang Huy – Ngô Đình Hy – Phan Thị Phương Uyên

21127428	Phạm Nguyễn Quốc Thanh
----------	------------------------



Giới thiệu đồ án:

Mục tiêu của đồ án là tìm hiểu các yếu tố quyết định mức lương và việc làm của các kỹ sư ngay sau khi tốt nghiệp. Các yếu tố như điểm số ở các cấp/trường đại học, kỹ năng của ứng viên, sự liên kết giữa trường đại học và các khu công nghiệp/công ty công nghệ, bằng cấp của sinh viên và điều kiện thị trường cho các ngành công nghiệp cụ thể sẽ ảnh hưởng đến điều này.

Bộ dữ liệu được sử dụng trong đồ án này thu thập tại Ấn Độ, nơi có hơn 6000 cơ sở đào tạo kỹ thuật công nghệ với khoảng 2,9 triệu sinh viên đang học tập. Mỗi năm, trung bình có 1,5 triệu sinh viên tốt nghiệp chuyên ngành Công nghệ/Kỹ thuật, tuy nhiên do thiếu kỹ năng cần thiết, ít hơn 20% trong số họ có việc làm phù hợp với chuyên môn của mình. Bộ dữ liệu này không chỉ giúp xây dựng công cụ dự đoán mức lương mà còn cung cấp thông tin về các yếu tố ảnh hưởng đến mức lương và chức danh công việc trên thị trường lao động.

Các thư viện sử dụng:

- [1] numpy
- [2] pandas
- [3] IPython
- [4] sklearn

Các hàm và lớp được sử dụng:

`def mae(y, y_hat)`: Tính độ đo MAE. Sử dụng công thức ở phía dưới để tính toán. Hàm `ravel` được dùng để đưa giá trị `y` và giá trị `y` mũ về mảng 1 chiều và `mean` được dùng để tính trung bình.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Lớp `OLSLinearRegression`: lớp được dùng để sử dụng mô hình hồi quy tuyến tính. Trong đây, có 3 hàm chính:

`def fit(self, X, y)`: Được sử dụng để tính toán các trọng số cho mô hình dựa trên giá trị mục tiêu. Sử dụng công thức ở phía dưới để tính toán. Trong đây, `X_pinv` tương đương với $(X^T X)^{-1} X^T$, hàm `linalg.inv` được dùng để lấy nghịch đảo của $(X^T X)$. Dấu “@” được dùng như một phép nhân giữa các ma trận.

$$w = (X^T X)^{-1} X^T y$$

`def get_params(self)`: Được sử dụng để lấy các trọng số `w` sau khi khớp dữ liệu cho mô hình.

`def predict(self, X)`: Được sử dụng để tính các giá trị mục tiêu mới có thể có sau khi khớp dữ liệu, sử dụng các trọng số đã được tính ở phía trên và một bộ dữ liệu `X`.

`def format_float(x)`: Được sử dụng để định dạng lại các giá trị MAE làm tròn đến 3 chữ số.

Kết quả và nhận xét:

a. Sử dụng 11 đặc trưng đầu tiên:

- Giải thích:

- o `lr_1a` là biến được dùng để gọi lớp `OLSLinearRegression` và gọi hàm `fit` để khớp dữ liệu giữa 11 đặc trưng đầu tiên và tập huấn luyện.
- o `y_hat` là các giá trị mục tiêu mới sau khi đã khớp mô hình dữ liệu sử dụng hàm `predict` ở lớp `OLSLinearRegression` với tham số `X` là `first_11_data_test`.
- o Gọi hàm `mae` để tính độ đo MAE giữa giá trị mục tiêu của tập kiểm tra và giá trị mục tiêu của mô hình.

- Kết quả:

Đặc trưng	Trọng số
Gender	-22756.513
10percentage	804.503
12percentage	1294.655
CollegeTier	-91781.898
Degree	23182.389
collegeGPA	1437.549
CollegeCityTier	-8570.662
English	147.858
Logical	152.888
Quant	117.222
Domain	34552.286
MAE	104863.778

- Công thức:

$$\begin{aligned}
 \text{Salary} = & -22756.513 \times \text{Gender} \\
 & + 804.503 \times 10\text{percentage} \\
 & + 1294.655 \times 12\text{percentage} \\
 & + -91781.898 \times \text{CollegeTier} \\
 & + 23182.389 \times \text{Degree} \\
 & + 1437.549 \times \text{collegeGPA} \\
 & + -8570.662 \times \text{CollegeCityTier} \\
 & + 147.858 \times \text{English} \\
 & + 152.888 \times \text{Logical} \\
 & + 117.222 \times \text{Quant} \\
 & + 34552.286 \times \text{Domain}
 \end{aligned}$$

- Nhận xét: Qua bảng trên, ta dễ dàng thấy các đặc trưng có trọng số âm là Gender, CollegeTier và CollegeCityTier, trong đó trọng số âm nhỏ nhất là CollegeTier. Các đặc trưng còn lại đều là dương và đặc trưng có trọng số lớn nhất là Degree. CollegeTier được quy định bởi điểm trung bình AMCAT đạt được bởi sinh viên tại trường đại học đó, các trường đại học có điểm trung bình vượt qua ngưỡng cho trước được đánh dấu là 1 và các trường khác là 2. CollegeCityTier là hạng của thành phố có trường đại học mà ứng viên theo học, được quy định dựa trên dân số của thành phố. Gender là giới tính của ứng viên. CollegeTier và CollegeCityTier cho ta thấy về sự khác biệt ở mức lương giữa các khu vực và các trường đại học. Các ứng viên tốt nghiệp trường đại học top 1 đều có lương tốt hơn so với các ứng viên

khác, và mức lương của những ứng viên ở các khu vực khác nhau đều sẽ có chênh lệch. Về phần giới tính, việc có trọng số âm thể hiện rằng có sự khác biệt về lương giữa nam và nữ, do ta đặt 2 là dành cho nữ và 1 là dành cho nam thế nên ta có thể thấy rằng lương của kỹ sư nam trong ngành sẽ cao hơn nữ đồng nghiệp. Về các trọng số dương, Degree có trọng số lớn nhất với Degree 1 là bằng cử nhân và 2 là bằng thạc sĩ. Điều này thể hiện rằng sự khác biệt của mức lương giữa các cử nhân và thạc sĩ là khá lớn. Do việc học thạc sĩ sẽ giúp cho ta phát triển chuyên sâu về 1 mảng, giúp ta có thêm các mối quan hệ với các cá nhân tài năng từ nhiều nhóm ngành khác nhau, rèn luyện khả năng tư duy, kỹ năng giải quyết vấn đề và sẽ được đảm bảo việc làm (ngay cả trong thời kỳ kinh tế không ổn định). Đối với các đặc trưng còn lại, đây đều là những đặc trưng liên quan tới học thuật và các kỹ năng cá nhân. Các đặc trưng học thuật bao gồm 10percentage, 12percentage và CollegeGPA, tất nhiên để trở thành 1 kỹ sư cũng cần phải học rất nhiều và tiếp thu rất nhiều kiến thức thế nên các đặc trưng này dương là chuyện tất nhiên. Các đặc trưng còn lại là English, Logical, Quant và Domain, ở đây, ta sẽ nói về Domain vì 3 đặc trưng còn lại sẽ được giải thích rõ hơn ở câu c. Domain ở đây là điểm số trong module chuyên ngành. Kiến thức chuyên ngành, dù ít được nói tới khi bàn luận về những gì cần thiết để thành công, nhưng lại rất quan trọng trong việc giúp chúng ta làm việc năng suất hơn. Các kỹ sư thường xuyên phải giải quyết các vấn đề rất phức tạp trong công việc của họ, sự hiểu biết sâu về lĩnh vực giúp họ xác định được các cơ hội để cải tiến và sáng tạo. Các kỹ sư thường làm việc cùng đồng nghiệp, khách hàng và các bên liên quan. Sở hữu kiến thức chuyên ngành giúp giao tiếp rõ ràng và hiệu quả. Khi mọi người đều sử dụng cùng một ngôn ngữ kỹ thuật, nguy cơ hiểu lầm được giảm thiểu và ý tưởng có thể được truyền đạt lại một cách chính xác. Các lĩnh vực kỹ thuật phát triển theo thời gian với các công nghệ, kỹ thuật và nghiên cứu mới. Các kỹ sư có kiến thức chuyên ngành tốt sẽ nắm bắt kịp thời để cập nhật những thay đổi này, thích nghi với những phát triển mới và duy trì vị trí hàng đầu trong lĩnh vực của họ. [3] [8] [9] [10] [14] [16]

b. Sử dụng các đặc trưng tính cách để tìm mô hình tốt nhất:

- Ý tưởng: Dựa vào [1], ta muốn xem xét rằng liệu tính cách nào là tốt nhất cho các kỹ sư. Có 5 tính cách chính: 'conscientiousness', 'agreeableness', 'extraversion', 'neuroticism' và 'openess_to_experience'. Dựa vào kỹ thuật k-fold cross validation để tính độ đo MAE tốt nhất cho mỗi đặc trưng. Với mỗi tính cách, ta đều chia ra k_fold bộ dữ liệu dựa trên bộ dữ liệu đã được xáo trộn. Thực hiện xáo trộn trên toàn bộ tập huấn luyện để mỗi bộ dữ liệu đều được khớp với giá trị mục tiêu đã có. Sử dụng 1 bộ dữ liệu để kiểm tra và k - 1 bộ còn lại để huấn luyện.
- Giải thích: train là bộ dữ liệu đã được xáo trộn. y_train được dùng để lưu lại các giá trị mục tiêu sau khi xáo trộn. models_train là danh sách bao gồm các mô hình sẽ được sử dụng. Ở đây là 'conscientiousness', 'agreeableness', 'extraversion', 'neuroticism' và 'openess_to_experience'. Tất cả mô hình đều được reshape lại thành 1 cột. models_test cũng tương tự nhưng là dành cho sau khi đã kiểm được mô hình tốt nhất. average_maes là danh sách các kết quả MAE của mỗi mô hình. Sử dụng hàm cross_val_score [] và mean để tính toán giá trị trung bình của mỗi mô hình trong k_fold lần. best_model_index là biến được dùng để lấy index của mô hình có giá trị MAE thấp nhất. df là bảng so sánh các mô hình và giá trị MAE.
- Kết quả:

STT	Mô hình với 1 đặc trưng	MAE
4	neuroticism	299248.592104948
1	agreeableness	300918.33048528567
3	openess_to_experience	303138.5600062765
0	conscientiousness	306115.22946836246
2	extraversion	306960.13672823523

- Nhận xét: Có thể thấy rằng, 'neuroticism' có kết quả MAE tốt nhất trong 5 đặc trưng. Để lý giải cho điều này, ta vốn biết rằng đặc thù công việc của 1 kỹ sư vốn rất áp lực. Về neuroticism, một người sở hữu chỉ số cao về tính nhạy cảm(neuroticism) có các đặc điểm như: lo lắng nhiều thứ, hay overthinking, thường xuyên cảm thấy buồn phiền, tâm trạng bồn chồn, bất an. Ngược lại, những người có chỉ số thấp lại có các đặc điểm như: cảm xúc ở mức ổn định, biết cách đối mặt với sự căng thẳng... Chính vì vậy, ta có thể thấy rằng neuroticism có thể phù hợp với kỹ sư và trọng số của neuroticism cũng là một số âm, (được thể hiện ở phần sau) cho thấy rằng chỉ số thấp sẽ ảnh hưởng tới mức lương nhiều hơn. Với 2 đặc trưng kế tiếp là agreeableness và openness_to_experience, "Openness to experience" có thể hiểu sơ là người có tính cách cởi mở. Các đặc điểm của tính cách này là: sự sáng tạo, trí tuệ,..., "agreeableness" có thể hiểu là người dễ hoà thuận, dễ tính. Tất nhiên, là 1 kỹ sư nên tính sáng tạo là điều cực kì cần thiết vì 1 kỹ sư mà không dám đổi mới hay phá bỏ những lĩnh vực khác sẽ không bao giờ đạt được thành tựu gì cả và họ cũng thường xuyên phải giao tiếp với những người khác với những tính cách khác nhau nên họ cần phải có khả năng giao tiếp và được lòng những người khác để công việc được thuận lợi hơn. Tuy nhiên, ta cũng có thể thấy rằng độ đo MAE của cả 5 đặc trưng đều khá lớn so với độ đo MAE ở câu a. Điều này có nghĩa là mặc dù tính cách có ảnh hưởng tới mức lương của kỹ sư nhưng lại không quá nhiều mà còn phụ thuộc vào những đặc trưng khác. Thế nên, mức lương của 1 kỹ sư không nên được tính toán bằng cách xem xét 1 trong các tính cách của họ mà cần phải kết hợp nhiều đặc trưng khác và cũng nên xét các tính cách khác của họ. [2] [4] [5] [6] [7]
- Mô hình tính cách tốt nhất: neuroticism.

- Sau khi tính toán và so sánh các đặc trưng tính cách, ta lấy ra index của đặc trưng tốt nhất bằng cách sử dụng hàm `index(min(average_maes))` và lưu vào biến `best_model_index`. `my_best_personality_feature_model` sẽ là mô hình sử dụng tính cách tốt nhất và khớp dữ liệu với bộ dữ liệu `y_train_1b` đã được xáo trộn. `X_Para` là bộ dữ liệu sẽ được sử dụng để tính toán độ đo MAE và được dựa vào `best_model_index` để chọn. `y_hat` là các giá trị mục tiêu mới sau khi đã khớp mô hình dữ liệu sử dụng hàm `predict` ở lớp `OLSLinearRegression` với tham số `X` là `X_Para`. Sau đây, ta gọi hàm `mae` để tính độ đo MAE giữa giá trị mục tiêu của tập kiểm tra và giá trị mục tiêu của mô hình.
- Kết quả:
- Nhận xét: Đặc trưng "neuroticism" có một hệ số âm khá lớn (-56546.304). Điều này cho ta thấy đặc trưng này có tác động mạnh mẽ đến dự đoán mức lương. Trọng số âm cho ta thấy rằng khi "neuroticism" tăng, mức lương sẽ giảm. Có thể thấy rằng đặc trưng "neuroticism" có ảnh hưởng mạnh đến mức lương của kỹ sư theo hướng ngược nhau. Cụ thể là, khi mức độ "neuroticism" tăng, mức lương sẽ giảm. Điều này có thể liên quan đến việc những người có điểm số "neuroticism" cao có thể có hiệu suất làm việc thấp hơn hoặc có khả năng gây ra các vấn đề tương tác trong môi trường làm việc. Giá trị MAE của mô hình này cũng là khá cao so với câu a, cho ta thấy rằng không nên chỉ dựa vào tính cách mà dự đoán mức lương cho 1 kỹ sư mà cần phải xem xét các yếu tố khác. [2] [4] [7]

Trọng số	Đặc trưng	MAE
-56546.304	neuroticism	291019.693226953

- c. Sử dụng các đặc trưng về khả năng cá nhân như logic, ngoại ngữ và định lượng:
 - Ý tưởng: Ta muốn xem xét rằng liệu khả năng cá nhân của mỗi người có ảnh hưởng nhiều tới mức lương của họ hay không. Cách thức làm hoàn toàn giống câu b, ta cũng sẽ sử dụng kỹ thuật k-fold cross validation.
 - Kết quả:

STT	Mô hình với 1 đặc trưng	MAE
0	Quant	118144.20890177856
1	Logical	120332.9243730387
2	English	121948.94944543508

- Nhận xét: Có thể thấy rằng, Quant (khả năng định lượng) có kết quả MAE tốt nhất trong 3 khả năng và cả 3 độ đo MAE chênh lệch với nhau không quá lớn và với câu a cũng không quá lớn. Quant (Quantitative) (khả năng định lượng) là khả năng của 1 người để xử lý và giải quyết dữ liệu "numerical" và "categorical". Các kỹ sư thường phải làm việc với các con số, đo lường, tính toán và phân tích dữ liệu. Khả năng định lượng là cần thiết cho các nhiệm vụ như thiết kế thí nghiệm, phân tích kết quả, mô hình hóa hiện tượng vật lý và dự đoán chính xác. Kỹ sư cần phải thoải mái với các khái niệm toán học để giải quyết các vấn đề phức tạp và tối ưu hóa thiết kế. Logical là khả năng suy luận dựa trên lý lẽ. Các dự án kỹ thuật thường bao gồm các vấn đề về hệ thống và logic. Các kỹ sư cần phải chia các vấn đề phức tạp thành các bước có thể dễ dàng quản lý được. Chính vì vậy, 2 đặc trưng này được xem là 1 yếu tố cần thiết cho 1 kỹ sư. Tiếp theo, về khả năng ngoại ngữ, ngôn ngữ thường được sử dụng bởi các kỹ sư để có thể trao đổi, giải thích các ý tưởng cho nhau là tiếng Anh. Hơn nữa, hầu hết các bài báo cáo học thuật đều được viết bằng tiếng Anh. Thứ hai, tiếng Anh còn giúp bản thân ta giao tiếp với những người đến từ những đất nước khác, hoặc những khách hàng tiềm năng, những người ít nhất cũng kì vọng rằng bản thân ta biết nói tiếng Anh một cách chuẩn xác và lưu loát. Chính vì vậy, đây chính là ngôn ngữ quan trọng nhất mà mọi kỹ sư đều cần phải biết. Thế nên, việc cả 3 khả năng này đều có độ đo MAE tương đối ổn là điều khá dễ hiểu. [1] [5] [6]

- Mô hình khả năng cá nhân tốt nhất: Quant

- o Cách thức thực hiện tương tự ở đặc trưng neuroticism.
- o Kết quả:

Trọng số	Đặc trưng	MAE
585.895	Quant	106819.57761989674

- o Nhận xét: Trọng số của đặc trưng Quant là số dương, có nghĩa là điểm Quant càng cao sẽ ảnh hưởng tới mức lương của 1 kỹ sư và độ đo MAE của Quant cũng không chênh lệch quá nhiều so với câu a.

d. Xây dựng mô hình riêng:

- [1] Tìm mô hình: Thực hiện k-fold cross validation trên 24 mô hình tổng cộng. Với 23 mô hình đầu tiên chỉ sử dụng một đặc trưng và 1 mô hình cuối cùng là tất cả các đặc trưng.

- a. Kết quả:

STT	Feature	MAE
23	All	110419.18251298685
9	Quant	118144.20890177856
1	10percentage	118852.56650941682
2	12percentage	120033.09735642346
8	Logical	120332.9243730387
5	collegeGPA	121531.9618816522
7	English	121948.94944543508
3	CollegeTier	133475.59757248283
4	Degree	137527.57145541575
0	Gender	150416.3373767708
11	ComputerProgramming	156016.3763422918
10	Domain	175539.14229203324
6	CollegeCityTier	251450.68067432227
12	ElectronicsAndSemicon	257427.22980746388
13	ComputerScience	268409.9844779946
16	TelecomEngg	290913.93216654425
14	MechanicalEngg	297415.8094724575
21	nueroticism	299248.592104948
19	agreeableness	300918.33048528567
15	ElectricalEngg	301204.47237528616
22	openess_to_experience	303138.5600062765
18	conscientiousness	306115.22946836246
17	CivilEngg	306116.1567246794
20	extraversion	306960.13672823523

- b. Nhận xét: Có thể thấy mô hình sử dụng tất cả các đặc trưng có độ đo MAE tốt nhất, khả năng định lượng (Quant) là đặc trưng tốt nhất nếu xét các mô hình chỉ sử dụng 1 đặc trưng và các đặc trưng tính cách đều thuộc vào nhóm dưới, nhóm có các độ đo MAE tệ nhất. Việc sử dụng kỹ thuật k-fold cross validation và sử dụng tất cả các đặc trưng cho mô hình dự đoán cho ra kết quả MAE tốt nhất cho ta thấy rằng các đặc trưng kể trên đều ảnh hưởng tới mức lương của kỹ sư.

[2] Đưa ra các mô hình:

1. Mô hình sử dụng 23 đặc trưng. Lý do: Vì ta đã biết rằng tất cả các đặc trưng dù ít hay nhiều đều sẽ ảnh hưởng tới mức lương.

2. Mô hình không sử dụng các đặc trưng ‘MechanicalEngg’, ‘CivilEngg’. Lý do: Theo như báo The New Indian Express, mặc dù các lĩnh vực như xây dựng và sản xuất đang phát triển ở Ấn Độ, nhu cầu về các kỹ sư xây dựng và cơ khí vẫn giảm hàng năm. Nhiều chuyên gia chỉ ra các nguyên nhân cho hiện tượng này như là: lương thấp, thiếu cơ hội việc làm ổn định và chương trình học lạc hậu không trang bị cho sinh viên những kỹ năng cần thiết cho ngành nghề. Một số chuyên gia khẳng định rằng thay vì thuê các kỹ sư tốt nghiệp có trình độ trung bình, các công ty thích tuyển dụng các kỹ sư có chứng chỉ, những người có nhiều kiến thức thực tế hơn và yêu cầu mức lương thấp hơn. Một lý do khác có thể là sinh viên ngày nay thích công việc văn phòng, cho phép họ làm việc tại văn phòng có điều hoà. [11]

3. Mô hình không sử dụng đặc trưng ‘Degree’. Lý do: Vì các tấm bằng “vô dụng”. Theo các cuộc phỏng vấn của Bloomberg tại Ấn Độ, hàng triệu bạn trẻ nước này đã bỏ đi học đại học vì mong muốn không chỉ thăng tiến trong sự nghiệp mà còn muốn đổi đời, gia tăng địa vị xã hội, để lấy vợ hơn. Thậm chí vì quá lo lắng cho tương lai mà nhiều bạn trẻ Ấn Độ còn trả tiền để học thêm văn bằng 2-3 nhằm gia tăng cơ hội xin việc làm. Vậy là vô số thanh thiếu niên Ấn Độ chui vào những lớp đại học bên trong các tòa chung cư nhỏ, hoặc thậm chí tại các địa điểm quầy hàng ngoài chợ để đi học. Trên các tuyến đường cao tốc Ấn Độ cũng ngập tràn nhiều băng rôn quảng cáo về những học viện, trường đại học hứa hẹn xin việc làm đầu ra cho sinh viên.^{[12] [13] [15] [16]}

[3] Kết quả:

STT	Mô hình	MAE
1	Mô hình 22 đặc trưng (Degree))	110575.53694310327
2	Mô hình 21 đặc trưng (Mech, Civil)	110613.59217930939
0	Mô hình 23 đặc trưng	110761.27348783193

[4] Nhận xét: Cả 3 mô hình đều chênh lệch với nhau không quá nhiều và mô hình sử dụng 22 đặc trưng bỏ Degree sẽ là mô hình tốt nhất. Như vậy, ta có thể thấy được rằng vấn đề về tấm bằng cử nhân hay thạc sĩ ở Ấn Độ quan trọng hơn đang là vấn đề khá quan trọng. Khi mà theo nhiều bài báo, lương trung bình của các thạc sĩ sẽ cao hơn cử nhân. Lý giải cho việc này, ở Ấn Độ có rất nhiều trường lớp “Đại học” dưới chuẩn với những giảng viên thiếu kinh nghiệm, giáo trình lạc hậu. Hậu quả là rất nhiều bạn sinh viên ra trường nhưng lại thiếu trình độ, cần đào tạo lại. Báo cáo của hãng SHL cho thấy chỉ 3,8% số cử nhân ngành kỹ sư tại Ấn Độ là có đủ trình độ để ứng tuyển vào những công việc liên quan đến phần mềm tại các startup.

[5] Huấn luyện lại mô hình:

Kết quả:

Đặc trưng	Trọng số
Gender	-23699.803
10percentage	875.683
12percentage	1171.61
CollegeTier	-81235.349
collegeGPA	1772.254
CollegeCityTier	-5421.824
English	154.006
Logical	123.131
Quant	101.7
Domain	28036.853
ComputerProgramming	77.934
ElectronicsAndSemicon	-49.342
ComputerScience	-178.44
MechanicalEngg	33.156
ElectricalEngg	-153.572
TelecomEngg	-64.832
CivilEngg	146.868
conscientiousness	-19799.095
agreeableness	15230.887
extraversion	5054.964
neuroticism	-10592.207
openess to experience	-5725.161
MAE	101582.134

Nhận xét: Như vậy, mô hình này có kết quả MAE tốt nhất trong các mô hình trên. Điều này cho ta thấy được rằng tám bằng ở Ấn Độ đang là 1 vấn đề cực kì nghiêm trọng, khoảng cách mức lương giữa giới tính là vấn đề cấp bách và sự khác biệt giữa các trường đại học top đầu và phần còn lại vẫn còn khá lớn.

Tài liệu tham khảo:

- [1] [Quantitative analysis of historical data for prediction of job salary in India - A case study](#)
- [2] [Here's How Your Personality Type May Affect Your Income](#)
- [3] [Your high school GPA could affect your income](#)
- [4] [Should I Become an Engineer?](#)
- [5] [The importance of using English Language effectively in Engineering](#)
- [6] [Do Engineers Need to be Fluent in English?](#)
- [7] [Big 5 Personality Là Gì? Tất Tần Tật Về Mô Hình Tính Cách 5 Yếu Tố](#)
- [8] [Does GPA matter for university graduates' wages?](#)
- [9] [Gender Pay Gap in India: A Reality and the Way Forward—An Empirical Approach Using Quantile Regression Technique](#)
- [10] [How Much Does a Degree Increase Your Earnings? \(With Median Salaries\)](#)
- [11] [Why are civil, mechanical engineering unwanted?](#)
- [12] [Chuyện không chỉ ở riêng Việt Nam: Đa phần kỹ sư Ấn Độ ra trường phải đi làm công nhân](#)
- [13] [Ấn Độ và nghịch lý nhân tài đông đảo nhưng lại có cả một thế hệ thất nghiệp vì những tám bằng 'vô dụng'](#)
- [14] [Coding is Not Enough: The Importance of Investing in Domain Knowledge](#)

-
- [15] [Án Độ hơn 90% thạc sĩ thất nghiệp](#)
[16] [Tại sao nên học thạc sĩ?](#)