













**Lời cảm ơn**

Lời đầu tiên em xin được gửi lời cảm ơn sâu sắc tới thầy cô giáo trong khoa Công Nghệ Thông tin đã dìu dắt, giúp đỡ, hướng dẫn em trong suốt khoảng thời gian em học tập và làm báo cáo thực tập ở trường.

Em xin được gửi lời cảm ơn chân thành tới Ban giám hiệu trường Đại Học Văn Lang, Ban chủ nhiệm Khoa Công Nghệ Thông Tin đã cho em một môi trường học tập và rèn luyện bản thân thật tốt. Em cảm ơn đã cho em cảm nhận được không khí của đại gia đình Văn Lang từ những ngày đầu bước chân vào ngưỡng cửa đại học và bắt đầu quá trình trưởng thành tại trường.

Em cũng xin bày tỏ lòng biết ơn sâu sắc tới chị Trần Hồ Trúc Thụy công ty TNHH CUỘC SỐNG VUI KHỎE đã cho em cơ hội được trải nghiệm và phát triển bản thân khi thực tập tại công ty. Chị rất nhiệt tình chào đón và hỗ trợ em thực hiện đề tài một cách tận tình. Và quan trọng hơn chị đóng vai trò cung cấp tài liệu và chia sẻ một khối lượng không nhỏ những kiến thức thực tế không chỉ về mặt công nghệ mà còn về mặt xã hội, đời sống, khoa học và con người.

Em xin gửi lời cảm ơn chân thành và sâu sắc tới thầy Phan Hồ Viết Trường đã hướng dẫn em hoàn thành khóa thực tập này.

Em xin trân trọng cảm ơn!

*TP. Hồ Chí Minh, ngày tháng năm* 2023

**Sinh viên thực hiện**

*(Ký và ghi rõ họ tên) (Ký và ghi rõ họ tên)*

**CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**Độc lập – Tự do – Hạnh phúc**

**PHIẾU NHẬN XÉT SINH VIÊN**

**CỦA ĐƠN VỊ THỰC TẬP**

**Họ và tên sinh viên**: Nguyễn Thành Đạt

**Lớp**: CNTT22 **Khóa**: K27 **Khoa**: Công nghệ thông tin

**Trực thuộc Trường**: Đại học Văn Lang

Trong thời gian từ ngày tháng năm 2023 đến ngày 31 tháng 5 năm 2023

**Tại Công ty**: Công ty TNHH Cuộc sống Vui khỏe

**Địa chỉ**: 82 Phan Đăng Lưu, Phường 5, Phú Nhuận, TP.Hồ Chí Minh

Sau quá trình thực tập tại đơn vị của sinh viên, chúng tôi có một số nhận xét đánh giá như sau:

| **STT** | **Mục nhận xét** | **Nội dung nhận xét** | **Điểm**  **(thang điểm 10)** |
| --- | --- | --- | --- |
| 1 | ***Về ý thức tổ chức kỷ luật, thái độ*** | - Tuân thủ các quy định của công ty  - Có trách nhiệm với công việc được giao |  |
| 2 | ***Về kỹ năng (giao tiếp, quan hệ, học tập, làm việc)*** | - Giao tiếp tốt  - Có sự gắn kết các thành viên trong tổ  - Có tinh thần tự học hỏi và hoàn thành công việc đúng hạn |  |

***Đánh giá chung sau khi thực tập:***

Ngày tháng năm 2023

**Người hướng dẫn trực tiếp Xác nhận của đơn vị thực tập**

*(Ký và ghi rõ họ tên) (Ký và ghi rõ họ tên, và đóng mộc Cty)*

**Trần Hồ Trúc Thụy**

**CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**Độc lập – Tự do – Hạnh phúc**

**PHIẾU NHẬN XÉT SINH VIÊN**

**CỦA GIẢNG VIÊN HƯỚNG DẪN**

1. **Họ và tên sinh viên:** Nguyễn Thành Đạt

**MSSV**: 2174802010644 **Lớp**: K27CNTT22 **Khoa:** Công nghệ thông tin

**Đơn vị thực tập**: Công ty TNHH Cuộc sống Vui khỏe

**Giảng viên hướng dẫn**: ThS. Phan Hồ Viết Trường

1. **Nhận xét**:

*TP. Hồ Chí minh, ngày tháng năm 2023*

**Giảng viên hướng dẫn**

*(Ký và ghi rõ họ và tên)*

**ThS. Phan Hồ Viết Trường**

**Mục Lục**

[CHƯƠNG 1: MỞ ĐẦU 8](#_4d34og8)

[1.1 Lý do chọn chủ đề nghiên cứu: 8](#_2s8eyo1)

[CHƯƠNG 2: TÌM HIỂU CÔNG TY TNHH CUỘC SỐNG VUI KHỎE 8](#_17dp8vu)

[2.1 Giới thiệu lịch sử hình thành & quá trình phát triển của tập đoàn: 8](#_3rdcrjn)

[CHƯƠNG 3: TỔNG QUAN VỀ PHÂN TÍCH DỰ BÁO VỚI DATAIKU 8](#_26in1rg)

[3.1 Giới thiệu Flask 8](#_lnxbz9)

[3.2 Nguyên lý hoạt động của Flask 8`](#_35nkun2)

[3.3 Ví dụ minh họa Flask cơ bản: 8](#_1ksv4uv)

[3.4 Giới thiệu mô hình Random Forest: 8](#_44sinio)

[3.5 Kiến trúc mô hình Random Forest: 8](#_2jxsxqh)

[3.6 Ứng dụng của Random Forest trong thực tiễn 8](#_z337ya)

[CHƯƠNG 4: ÁP DỤNG MÔ HÌNH RANDOM FOREST ĐỂ DỰ BÁO MỨC ĐỘ DINH DƯỠNG CỦA MỘT MÓN ĂN 8](#_3j2qqm3)

[4.1 Giới thiệu dinh dưỡng trong bữa ăn: 8](#_1y810tw)

[4.2 Lợi ích dinh dưỡng đối với con người 8](#_4i7ojhp)

[4.3 Phát biểu vấn đề 8](#_2xcytpi)

[4.4 Thu thập dữ liệu 8](#_1ci93xb)

[4.5 Thiết kế Model 8](#_3whwml4)

[4.6 Thiết kế giao diện 8](#_2bn6wsx)

[4.7 Kết quả chương trình 8](#_qsh70q)

[CHƯƠNG 5: KẾT LUẬN VÀ ĐỀ XUẤT 8](#_3as4poj)

[5.1 Kết quả thực tập: 9](#_1pxezwc)

[5.2 Các kiến nghị rút ra từ kết quả của đợt thực tập: 9](#_49x2ik5)

[KẾ HOẠCH THỰC TẬP 10](#_2p2csry)

[TÀI LIỆU THAM KHẢO 12](#_147n2zr)

[NHẬT KÝ HƯỚNG DẪN BÁO CÁO THỰC TẬP 14](#_3o7alnk)

**Mục lục hình ảnh**

*[Image 2.1 Logo TMAI SaiGon 12](#_ihv636)*

*[Image 2.2 Trang thông tin tổng hợp 14](#_32hioqz)*

*[Image 2.3 Trang cung cấp sản phẩm về Lifestyle 14](#_1hmsyys)*

*[Image 2.4 Trang cung cấp thiết kế bao bì 15](#_41mghml)*

*[Image 2.5 Trang thương mại điện tử nông sản 15](#_2grqrue)*

*[Image 3.6 Giao diện máy ảo 22](#_vx1227)*

*[Image 3.7 Cài đặt gói DSS\_dataiku 23](#_3fwokq0)*

*[Image 3.8 Giao diện máy ảo sau khi import gói DSS 23](#_1v1yuxt)*

*[Image 3.9 Giao diện sau khi chạy thành công 24](#_4f1mdlm)*

*[Image 3.10 Giao diện làm việc chính (DSS-Studio) 24](#_2u6wntf)*

*[Image 4.11 Calibration curve(Đường chuẩn) 25](#_19c6y18)*

*[Image 4.12 Lift charts 26](#_3tbugp1)*

*[Image 4.13 ROC Curve (Đường cong ROC) 27](#_28h4qwu)*

*[Image 4.14 Density chart (Biểu đồ mật độ) 28](#_nmf14n)*

*[Image 4.15 Ma trận hỗn loạn 29](#_37m2jsg)*

*[Image 4.16 Dữ liệu dự báo 30](#_1mrcu09)*

**CHƯƠNG 1: MỞ ĐẦU**

**1.1 Lý do chọn chủ đề nghiên cứu:**

**CHƯƠNG 2: TÌM HIỂU CÔNG TY TNHH CUỘC SỐNG VUI KHỎE**

**2.1 Giới thiệu lịch sử hình thành & quá trình phát triển của tập đoàn:**

**CHƯƠNG 3: TỔNG QUAN VỀ PHÂN TÍCH DỰ BÁO VỚI DATAIKU**

**3.1 Giới thiệu Flask**

**3.2 Nguyên lý hoạt động của Flask**

**3.3 Ví dụ minh họa Flask cơ bản:**

**3.4 Giới thiệu mô hình Random Forest:**

**3.5 Kiến trúc mô hình Random Forest:**

**3.6 Ứng dụng của Random Forest trong thực tiễn**

**CHƯƠNG 4:   
ÁP DỤNG MÔ HÌNH RANDOM FOREST ĐỂ DỰ BÁO MỨC ĐỘ DINH DƯỠNG CỦA MỘT MÓN ĂN**

**4.1 Giới thiệu dinh dưỡng trong bữa ăn:**

**4.2 Lợi ích dinh dưỡng đối với con người**

**4.3 Phát biểu vấn đề**

**4.4 Thu thập dữ liệu**

**4.5 Thiết kế Model**

**4.6 Thiết kế giao diện**

**4.7 Kết quả chương trình**

**CHƯƠNG 5: KẾT LUẬN VÀ ĐỀ XUẤT**

## **5.1 Kết quả thực tập:**

Sau khoảng thời gian thực tập, tôi nghiên cứu đã được kết quả như sau:

- Hiểu thêm được phân tích dự báo và một trong những bước để phân tích dự báo nhu cầu của khách hàng.

- Phân tích, xử lý dữ liệu và trực quan hóa bằng các charts trong mô hình dự báo.

- Vận dụng cái kỹ năng kiến thức đã học ở trường để giải quyết công việc, tuy nhiên em đã học hỏi thêm từ các anh chị từ đó mà tiến bộ hơn

- Trao dòi thêm kỹ năng tự học qua việc đọc tài liệu, sách và các bài viết về Phân tích dự báo (predictive analytics) từ các nguồn khách nhau để mở rộng thêm kiến thức và tư duy.

Tuy nhiên, thời gian có hạn nên còn nhiều kiến thức về Phân tích dự báo và một số công việc vẫn chưa hoàn thành kịp.

## **5.2 Các kiến nghị rút ra từ kết quả của đợt thực tập:**

Đợt thực tập này đã cho em một trải nghiệm một phần thiết thực về phân tích dự báo và những kiến thức hoàn toàn mới. Tuy bước đầu thực hiện còn nhiều khó khăn, song nhờ sự giúp đỡ nhiệt tình của các cô và anh chị trong công ty, em đã có thể hoàn thành đề tài này.

Thời gian thực tập còn ngắn nên em vẫn chưa học hỏi được nhiều những nghiệp vụ cũng như tiếp cận nhiều hơn với các công cụ các mô hình dự báo. Hy vọng đợt thực tập sắp tới sẽ có thời gian nhiều hơn để nhóm có thể có nhiều cơ hội trải nghiệp hơn ở doanh nghiệp.

Trong tương lai, em sẽ cố gắng học hỏi và tìm hiểu các kiến thức liên quan đến phân tích dự báo, để có thể ứng dụng vào công việc và ngày càng hoàn thiện hơn trong lĩnh vực này.

**KẾ HOẠCH THỰC TẬP**

| **Thời gian** | **Nội dung công việc** | | |
| --- | --- | --- | --- |
| **Sinh viên** | **Giảng viên hướng dẫn tại Khoa** | **Người hướng dẫn tại doanh nghiệp** |
| **Tuần đăng ký** | * Đăng ký làm Đồ án thực tập 2 | * Gửi thông tin đăng ký thực tập * Xét tiêu chuẩn làm Đồ án thực tập 2 của sinh viên đăng ký. * Phân công GVHD. * Công bố danh sách làm Đồ án thực tập 2, công ty thực tập và GVHD | * Cung cấp các yêu cầu cho chương trình thực tập tại doanh nghiệp |
| **Tuần chuẩn bị** | * Liên lạc với GVHD | * Hướng dẫn quy trình làm thực tập (công việc, deadline) * Lên kế hoạch báo cáo với GVHD. * Cung cấp các biểu mẫu thực tập. |  |
| **Tuần thứ nhất** | * Liên lạc với doanh nghiệp. * Hoàn tất các thủ tục theo yêu cầu của đơn vị nhận thực tập. * Sinh viên đến làm việc thực tế tại đơn vị thực tập và tìm hiểu doanh nghiệp. * Lập kế hoạch thực tập. | * Hướng dẫn SV chọn đề tài * Hướng dẫn SV lập kế hoạch thực tập. |  |
| **Tuần thứ 2** | * Sinh viên làm việc thực tế tại đơn vị thực tập và tìm hiểu các công việc thực hiện. * Chọn đề tài/lĩnh vực mình muốn nghiên cứu/thực tập (có cân nhắc đến thời gian hoàn thành và phạm vi nghiên cứu phù hợp với yêu cầu của học phần). * Viết và nộp đề cương chi tiết cho GVHD. | * Kiểm tra tiến độ thực hiện của sinh viên. * Hướng dẫn viết đề cương chi tiết. * Chỉnh sửa và thông qua đề cương chi tiết. |  |
| **Tuần thứ 3-7** | * Sinh viên làm việc thực tế tại đơn vị thực tập và phải thực hiện các công việc mà đơn vị thực tập giao * Sinh viên phải quan sát, tìm hiểu các hoạt động thực tế có liên quan chủ đề thực tập; phỏng vấn trực tiếp những người có liên quan; thu thập thông tin và dữ liệu chuẩn bị cho việc viết báo cáo Đồ án thực tập 2. | * Kiểm tra tiến độ thực hiện của sinh viên * Trao đổi và hướng dẫn sinh viên các vấn đề liên quan đến thực tập và viết báo cáo Đồ án thực tập 2. * Nhận Đề cương chi tiết của sinh viên và ký tên xác nhận đã duyệt. |  |
| **Tuần thứ 8** | * Sinh viên làm việc thực tế tại đơn vị thực tập. * Hoàn thiện và chỉnh sửa báo cáo Đồ án thực tập 2. | * Kiểm tra tiến độ thực hiện của sinh viên. * Kiểm tra nội dung báo cáo Đồ án thực tập 2 |  |
| **Tuần thứ 9** | * Sinh viên làm việc thực tế tại đơn vị thực tập. * Gửi báo cáo Đồ án thực tập 2 đến cơ quan thực tập xin xác nhận. * Nộp báo cáo Đồ án thực tập 2 cho khoa. | * Kiểm tra tiến độ thực hiện của sinh viên. | * Nhận xét và đánh giá báo cáo Đồ án thực tập 2 của sinh viên. |
| **Tuần thứ 10** | * Trình bày và báo cáo Đồ án thực tập 2 | * Chấm điểm báo cáo Đồ án thực tập 2 qua buổi báo cáo. |  |
| **Tuần thứ 11** | * Xem điểm tổng kết và phản hồi nếu có. | * Tổng hợp và công bố điểm cho sinh viên * Gửi email cảm ơn doanh nghiệp. |  |

*TP. Hồ Chí Minh, ngày …. tháng …. năm 2023*

| **Giảng viên hướng dẫn**  *(Ký và ghi rõ họ tên)* | **Sinh viên thực hiện (nhóm trưởng)** *(Ký và ghi rõ họ tên)* |
| --- | --- |

**TÀI LIỆU THAM KHẢO**

[1] TechFlow, “Dataiku trong AI”, 2022  
link:<https://techflow.vn/tech-data-hop-tac-voi-dataiku-de-thuc-day-ung-dung-tri-tue-nhan-tao-ai-trong-doanh-nghiep/>

[2] Cision, “Dataiku phân tích dữ liệu”, 2021  
link: <https://vn.prnasia.com/story/60613-11.shtml>

[3] VietNambiz, “Hệ hỗ trợ quyết định trong Dataiku”, 2019

Link: <https://vietnambiz.vn/he-ho-tro-quyet-dinh-decision-support-system-la-gi-20190810192517123.htm>

[4] MeeyLand, “Dataiku là gì?”, 2022   
link: <https://meeyland.com/dau-tu/dss-la-gi-nhung-loi-ich-ma-dss-dem-den-cho-doanh-nghiep/>

1. BigData Solution, “Tổng quan về Preditive analytics” truy cập ngày 6/11/2022, link:<https://bigdatauni.com/tin-tuc/tong-quan-ve-predictive-analytics-phan-tich-du-bao-phan-1-khai-niem-va-loi-ich.html>
2. devoteam, “What is Dataiku ?”, truy cập ngày 6/11/2022, link: <https://www.devoteam.com/expert-view/what-is-dataiku/>
3. VIBLO, “Confusion Matrix/Ma trân nhầm lẫn/Ma trân lỗi <P1>, truy cập ngày 7/11/2022 link: <https://viblo.asia/p/confusion-matrix-ma-tran-nham-lan-ma-tran-loi-p1-V3m5WQB7ZO7>
4. Dataiku Knowledge, “Model Evaluation”, truy cập ngày 8/11/2022, link: <https://knowledge.dataiku.com/latest/courses/intro-to-ml/predictive-modeling/model-eval-summary.html?highlight=roc%20curve>
5. GIM, “Understanding the AUC-ROC Curve in Machine Learning Classification”, truy cập ngày 13/11/2022, link: <https://analyticsindiamag.com/understanding-the-auc-roc-curve-in-machine-learning-classification/>
6. Microsolf, “Lift Chart (Analysis Services - Data Mining)", tuy cập ngày 17/11/2022,

link: <https://learn.microsoft.com/en-us/analysis-services/data-mining/lift-chart-analysis-services-data-mining?view=asallproducts-allversions>

1. Course, “Calibration Curves”, truy câph ngày 19/11/2022,

link: <https://www.geeksforgeeks.org/calibration-curves/>

1. CESTI, “Trực quan hóa dữ liệu – Phần 5: Các dạng biểu đồ thể hiện sự phân phối của dữ liệu” , truy cập ngày 21/11/2022,

Link: http://thongke.cesti.gov.vn/dich-vu-thong-ke/tai-lieu-phan-tich-thong-ke/963-truc-quan-hoa-du-lieu-phan-5-cac-dang-bieu-do-the-hien-su-phan-phoi-cua-du-lieu

**TRƯỜNG ĐẠI HỌC VĂN LANG**

**KHOA CÔNG NGHỆ THÔNG TIN**

# **NHẬT KÝ HƯỚNG DẪN BÁO CÁO THỰC TẬP**

1. **Tên đề tài**: Tìm hiểu công cụ Dataiku trong phân tích dự báo thu nhập của khách hàng
2. **Giảng viên hướng dẫn**: ThS. Phan Hồ Viết Trường
3. **Sinh viên :** Đoàn Nguyễn Nhật Linh
4. **MSSV:** 197CT22329. **Lớp:** K25 – IT04
5. **Sinh viên :** Kim Hương Nam
6. **MSSV:** 197CT09864. **Lớp:** K25 – IT08
7. **Ngành:** Công nghệ thông tin

| **Thời gian** | **Nội dung làm việc** | **Nhận xét của GVHD**  *(Tiến độ, thái độ)* | **Các vấn đề chỉnh sửa** | **Xác nhận đã chỉnh sửa các nội dung theo yêu cầu của GVHD** | **Chữ ký của GVHD** | **Chữ ký của SV thực hiện** |
| --- | --- | --- | --- | --- | --- | --- |
|  | Tìm hiểu về phân tích dự báo |  |  |  |  |  |
|  | Cài đặt máy ảo VirtualBox và DSS\_Dataiku |  |  |  |  |  |
|  | Nhận Dataset từ công ty |  |  |  |  |  |
|  | Nhúng Dataset vào Dataiku và làm sạch |  |  |  |  |  |
|  | Đưa dataset vào Mô hình Machine learning Basics |  |  |  |  |  |
|  | Tim hiểu và đưa mô hình vào phân tích trực quan hóa |  |  |  |  |  |
|  | Tìm hiểu cơ bản về các chart trong mô hình. Xuất dữ liệu dự báo |  |  |  |  |  |
|  | Tìm hiểu về ưu nhược điểm của chart |  |  |  |  |  |

*TP. Hồ Chí minh, ngày tháng năm 2023*

**Giảng viên hướng dẫn**

*(Ký và ghi rõ họ và tên)*

**ThS.Phan Hồ Viết Trường**