**Employee Attrition Model Using R Learning Plan**

**1. Goal**

Build an Employee Attrition Model Using R on IBM HR Analytics Dataset with accuracy > 80% by August 19

**2. Resources**

*Main*

**1.** [**IBM HR Analytics Employee Attrition & Performance**](https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset): This is the perfect project as it'll force me to play with a dataset in my newly found R programming skills. This'll also reinforce my approach towards building machine learning models while making the project relevant to my job in people analytics in the future.

*Supplementary*

**2.** [**Udacity's Data Analysis with R**](https://classroom.udacity.com/courses/ud651): This will now be a supplementary resource, since I need to limit the scope of my time due to a higher priority project. After going through parts of the course, I don't think it's 100% necessary for me to finish the course before building the model. I can just keep building the model.

**3. Introduction to Statistical Learning With Applications in R:** This textbook will be my go-to for information regarding Machine Learning Models.

**3. Timeline & Key Milestones**

1. First Attrition Model Built - *Due Date: August 1*

*Justification: Shouldn't take too long to build a dumb model. The only gap in knowledge is learning how to implement the ML packages, but given my abilities in Python, that shouldn't be a huge problem.*

3. Final Attrition Model Built - *Due Date: August 13*

*Justification: Giving myself a week to tune up the model should give me enough time to hit my 80% accuracy mark. This'll also give me 5 days of cushion in case things don't go according to plan.*

**Final Due Date: August 13, 2017** - *Reason for date: This is set by SFDF accountability group, rather than personally set.*

***Estimated Hours per Week:*** *6-8 hours/week or 1.5-2 hours per day at minimum*

**4. Weekly Working Schedule**

***Preliminary Schedule:***

**Monday - Fri**

|  |  |
| --- | --- |
| **Time** | **Task** |
| 7 |  |
| 8 |  |
| 9 |  |
| 10 |  |
| 11 |  |
| 12 |  |
| 1 |  |
| 2 |  |
| 3 |  |
| 4 | **Project: HR Employee Attribution Model** |
| 5 |
| 6 |  |
| 7 |  |
| 8 |  |
| 9 |  |
| 10 |  |

**Saturday - Sunday**

*Rest Day*

**5. Accountability**

*Will pay someone $1000 if I don't finish by my deadline and will require weekly updates on project to be sent every Friday.*

**6. Sticking Points & Risk Management**

**Waking Up Too Late:** Especially on Mondays, there's a likelihood I'll wake up too late from staying out the whole weekend. My fear is that by waking up early on a Monday, then I'll be too tired and have no mental energy to build out the models. Do I take lessened Monday productivity to keep the same schedule or do I sleep in and stay more rested on Monday? In terms of overall productivity, sticking to the same schedule will force me to devote the time to working. So don't buy into the self-talk of needing more energy. Better to be tired Monday and make progress, even if it's not optimal. If I actually still wake up too late, I'll re-plan my schedule for that day and may need to work into the night, preferable to not do that.

**Hangover:** I can foresee this and I only usually go out on Friday & Saturday nights. Therefore, being hungover on a workday is extremely unlikely.

**Difficulty:** Getting stuck on a certain bug that won't allow my model to run, making me want to stop the project. I think accountability already forces me to try and figure this out and not give up.

**Going on unexpected trip:** The only thing I can actually foresee is going to Hawaii. To combat this, if this happens, I need to re-set my schedule when I get to Hawaii. This means, knowing the exact timeslots where I'll be needed beforehand so I can plan my day around the required 4 hours to commit to this project. Right when I get there, find 4 hours in the day, interrupted that I can commit to this project.

**Household clean-up help:** Since my house is a mess right now, my dad might need me to stay at home for a few days to help move heavy things. Worker arrival times are extremely unpredictable, since they might come late. If this is the case, talk to Dad and ask him what approximate time he might need me. Anything outside the time range, I will head to the library.

**Running Out of Cash, Can't Afford Ferry:** This would suck. But if this is the case, I'll shift my workplace to the Alameda Library. This would actually add more time for me since I won't have to take the ferry and getting to the library is much closer.

**7. Project & Performance Metrics**

Project Metric = *Lines of Code Written*

Performance Metric = *Accuracy Score%*

**8. Work Progress Tracking**

|  |  |
| --- | --- |
| **Date** | **Metrics** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**8. Weekly Project Reviews**

*What progress have I made this week? How am I doing on my performance metrics?*

*Am I doing the very best I can to accomplish my project's goal?*

*How can I improve/adjust this project plan to better achieve my goal?*

7/20: Felt like the course is a bit too basic on the data/EDA side, as I know everything. But the implementation in R is what I don't know and am learning. I do think the course has some value in showing me the commands, but the value is more in learning R than the EDA. Felt like I could actually start my project in R, since I basically know the logic of how to do it in Python. I'll just need to translate that into R by going through Google. Might look to change this to half course half modeling instead of going full course then modeling.