CIS 400 Senior Project
Spring 2018 Final Report
Team 17: Sleep-Dx

1. Student Information

Samuel Lerner
Kanishka Rao
Zane Stiles
Michael Terracciano
Project Name: Sleep-Dx

2. Advisor Information

Our CIS faculty advisor is Professor James Weimer . We will also be advised by Dr. Nalaka Gooneratne of the Penn Sleep Center. We have had the opportunity to meet with both Professor Weimer and Dr. Gooneratne several times this semester.

Overall, we began the semester with the initial goal of combining the signal processor and classifier into a single client-facing platform that doctors could interface with. Early meetings centered around how to best integrate with the clinical workstream and how to improve the output of our classifier. We then focused in-depth on the UI in preparation for the M&T Summit and CIS Senior Design Fair, and completed statistical and medical evaluations with both advisors.

We met with Dr. Gooneratne Feb 1 at 5 p.m. where we discussed setting up meetings with stakeholders and heard his own suggested user stories. Dr. Gooneratne suggested we develop our product as a plugin for Sandman, the program we observed sleep researchers using last semester. Our product would risk stratify for cardiopulmonary disease and fit into their existing workflow easily. We also began discussing the business opportunity for our product as part of our M&T requirements.

We met with Dr. James Weimer on Feb 2 at 1 p.m. and discussed testing the robustness of the classifier through a sensitivity analysis. Dr. Weimer suggested specific ensemble methods to improve our classifier, and explained what would be necessary to take our software from a working classifier to a fully-packaged enterprise product.

We Skype called Dr. Gooneratne on February 15 at 9:30 a.m. where we discussed in depth the market for sleep products, competitors, and the differences between customers and end-users. We are using our newfound knowledge of the Healthcare system to better approach our business analysis and understand our value proposition.

We discussed the business analysis with Dr. Gooneratne on February 20 and our M&T Advisor on February 23rd. They both provided relevant feedback regarding what details were most important and how to balance explaining technical and business concepts.

We spoke with professor Weimer on March 16 to discuss our progress on the platform and how to develop demo functionality in time for the M&T Summit

At the M&T Summit on March 24, we spoke with both our advisors in-depth about what aspects of our platform we should improve for the final sprint and how to re-prioritize with an end date in mind.

On April 23 we evaluated our clinical platform with Dr. Gooneratne and got his feedback on the platform as a proof of concept and discussed what next steps would be to take the product to market.

We have one final meeting planned with professor Weimer to handoff all the code and explain how to leverage it for future research.

3. Summary

We developed a data-analysis platform with integrated GUI to provide clinicians with probabilistic diagnoses for non-sleep-related conditions (specifically COPD) from sleep study data that can achieve 75% precision and recall and save the US Health System ~$350M/yr.

4. Overview of Problem and Approach

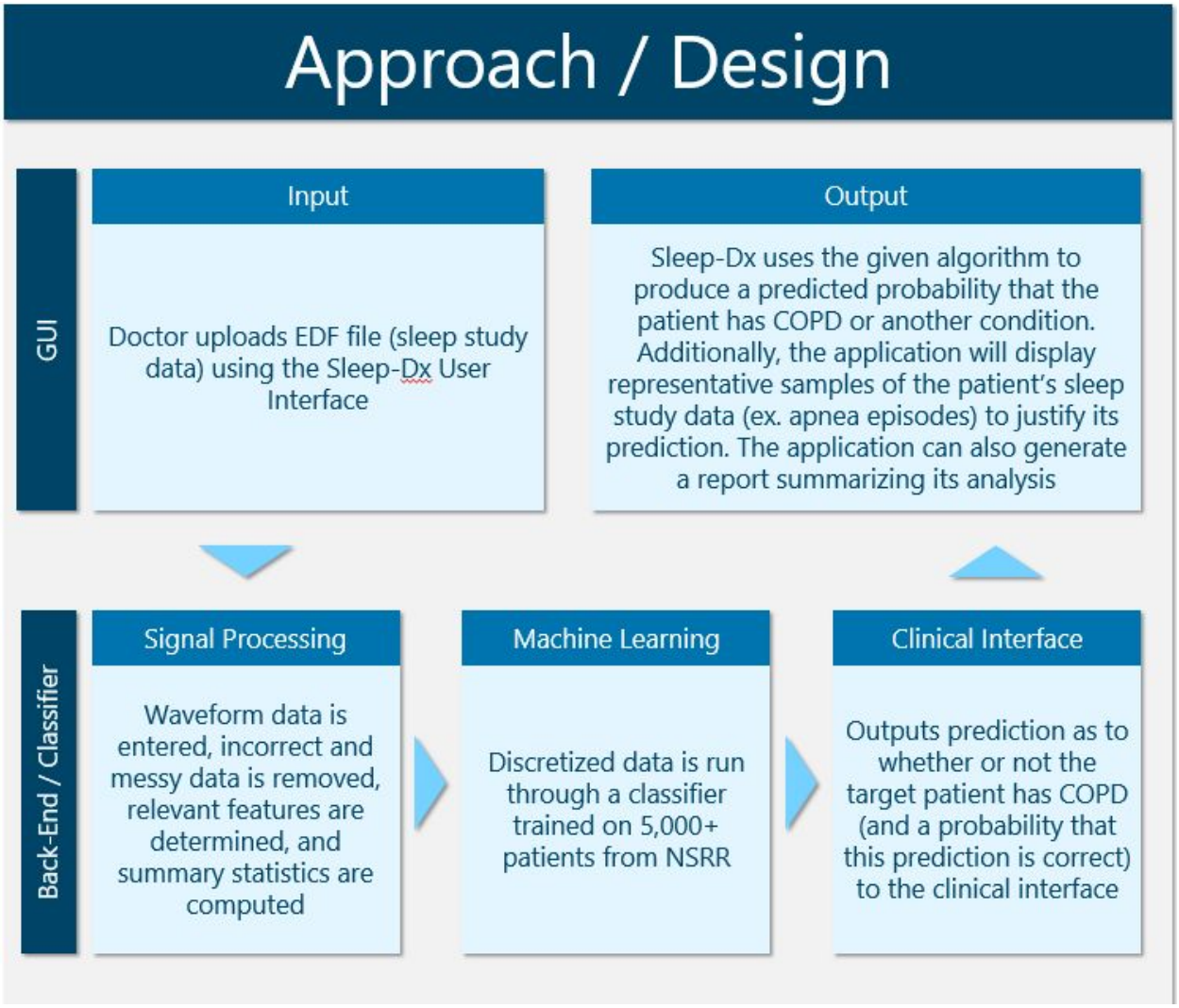
Sleep study results are a rich source of health data, but are currently underutilized as a method of diagnosing non-sleep-related conditions (e.g. depression, asthma, emphysema, etc.) As a result, patients may needlessly go undiagnosed for severe conditions after a sleep study and/or have to undergo tests other than a sleep study to diagnose these non-sleep-related conditions. We want to alleviate this problem by creating a classifier that can take in sleep study data and output a diagnosis for non-sleep related conditions (specifically COPD). This problem is difficult to solve, in part, because developing a classifier requires an analysis of waveform biometric data, which is continuous as opposed to discrete.

To solve this problem, we first needed to understand the data we dealing with and what the true predictors of COPD were. We then needed to obtain the relevant data in order to train a classifier to predict with high accuracy. Before training however, we developed a signal processor that abstracted and discretized the relevant data out of the waveform sleep data file, as the classifiers we were familiar with could not operate on a waveform datafile. We then implemented various different classifiers in order to optimize accuracy and arrive at a final classifier best equipped to diagnose new patients. Finally, we had to develop a clinician-facing platform that doctors could use to interface with the application and get the relevant data they need out of it.

5. Implementation

Architecture:

On the whole, our Architecture involves a front-end GUI with an input and results screen, with the signal processor working in the background. A simplified architecture can be seen below:

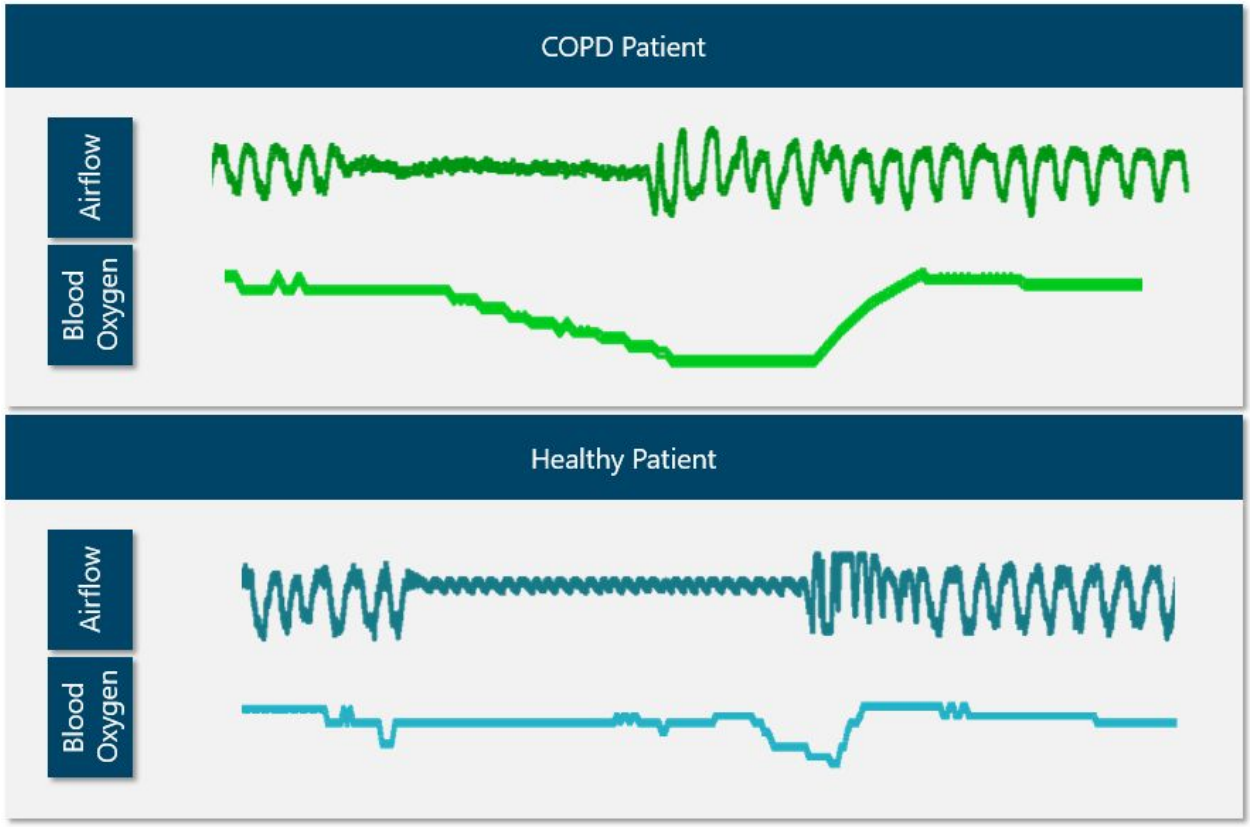


The goal of this modular architecture was two-fold, first to allow for easy integration of different parts and a structured development workflow, but also for future extensions on our platform, where clinical and technical researchers would be able to plug in various classifiers and processors and efficiently test their algorithms on sleep data. With the field of machine learning changing so rapidly, researchers better than us can plug-in better algorithms to any part of our application in order to impact more patients. Similarly, doctors can identify the relevant data in sleep studies to begin research on another condition and experiment with real data with all the setup abstracted away.

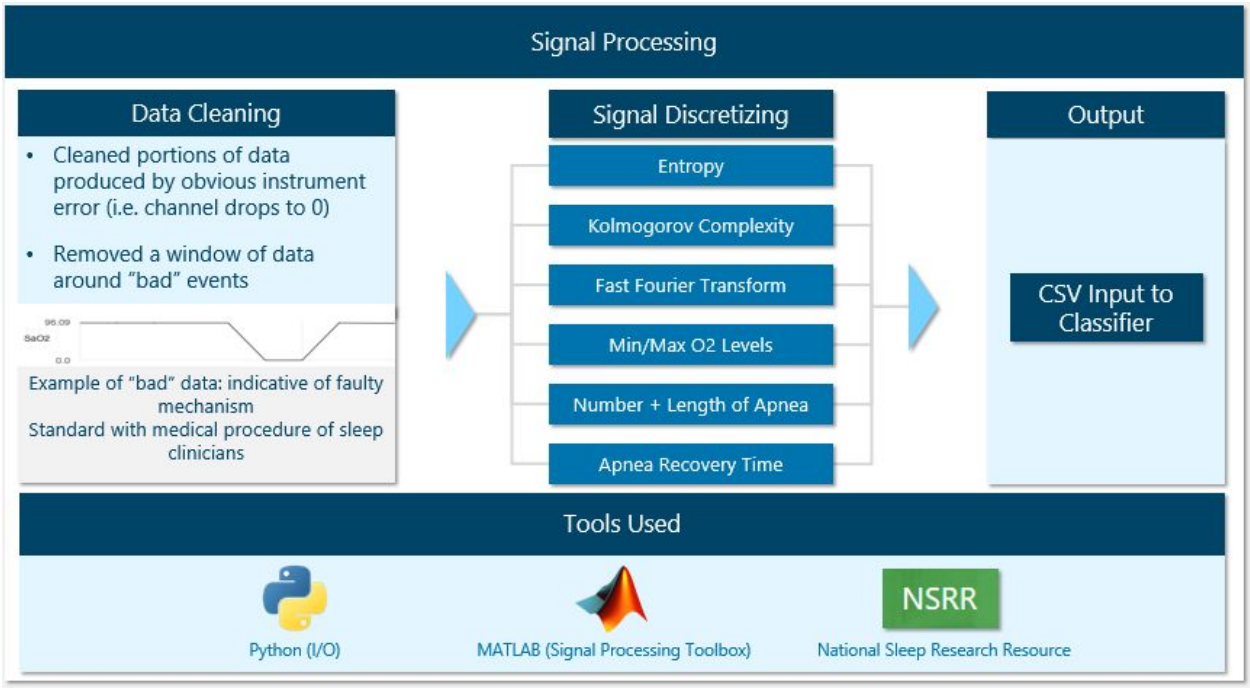
Signal Processor:

The three purposes of the signal processor were to remove 'bad' data, identify the relevant channels and segments of data, and compute discrete summary statistics that can be piped into a Machine Learning classifier. In terms of removing 'bad' data, when working with real-world medical data, there is significant 'messy' data generated from human and instrument error. An example is if your SO_2 levels drop to zero, this would mean you are dead. However, we actually see events like this occur quite often due to sensors falling off due to movement or poor attachment. Thus we programmed our signal processor

in Matlab to remove data sections where the data was ‘bad’. Next in terms of identifying relevant channels, most of this is hard coded as we spoke extensively with doctors at the Penn Sleep Center to identify what data would be most predictive for non-sleep related conditions. For COPD in particular, Oxymetry and Thorasic data was relevant, and the curves around apnea events are the segments we wanted to isolate. This leads into the third goal of computing the relevant statistics for a classifier. Take for example the below graph that demonstrates what occurs in a patient with COPD and one without:



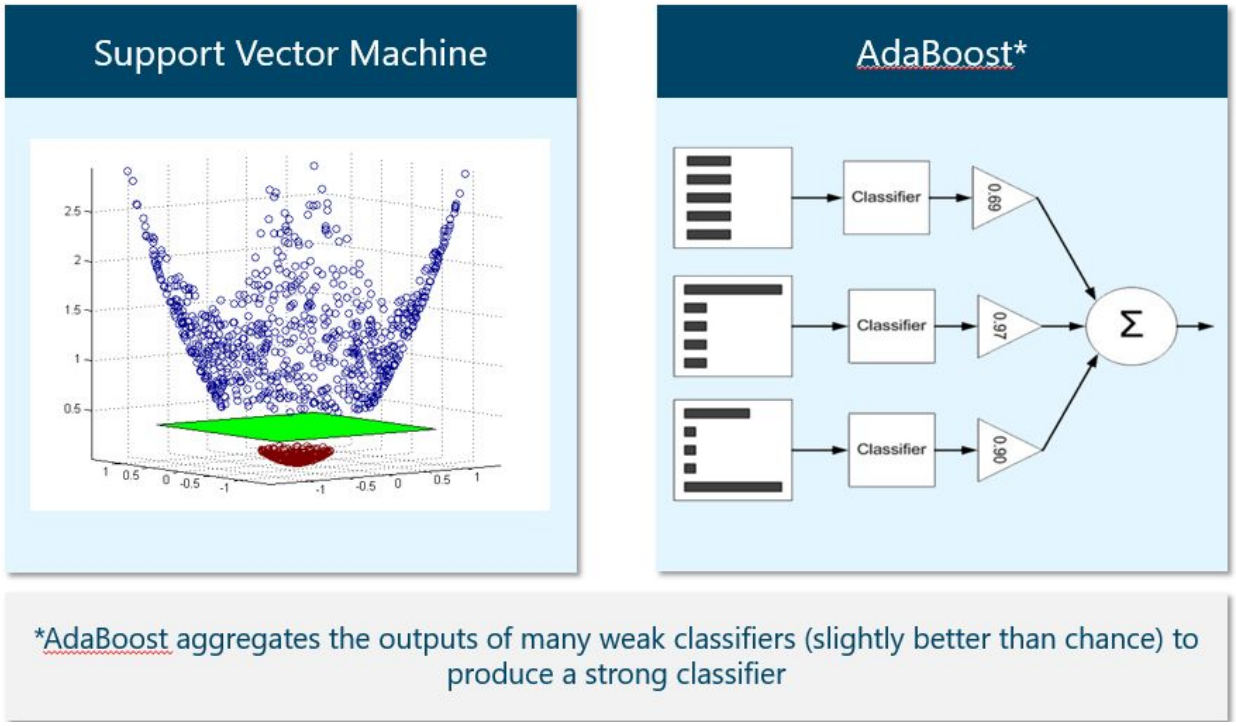
As you can see, during an apnea event, both SO_2 values drop, but the COPD patient has theirs drop more and take longer to recover. We extracted these events by capturing various statistics including frequency and duration of apnea events, the percentage drop in SO_2 values, and other statistical measures such as entropy. A summary slide for our signal processor can be found below:



Classifier:

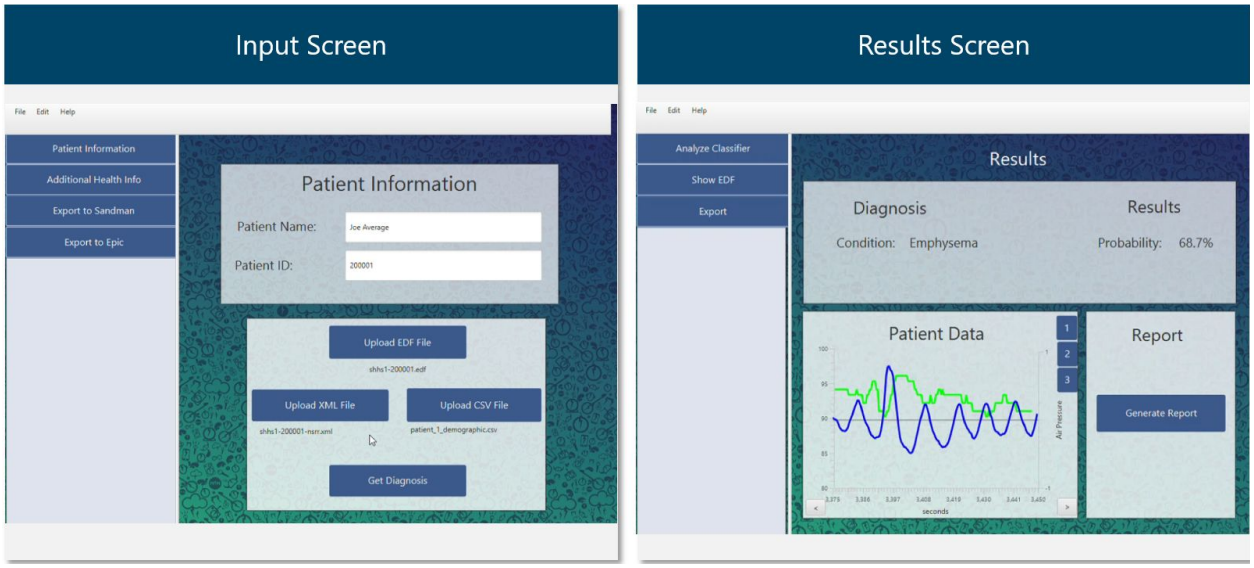
These calculated statistics are then fed into a Machine Learning classifier that predicts COPD and outputs the confidence percentage. Coming in with no knowledge, we arrived at a Support Vector Machine with Stochastic Gradient Descent as our final classifier after much trial and error. The classifier is a linear one that uses a 'hill-climbing' algorithm to try and find maximums to reweight different entries. We split our 5,000 data points in to roughly 80% training and 20% test during development as this is what we observed in similar experiments in the literature. Additionally, it was important that we used a larger training set as COPD is quite uncommon and running a smaller test set greatly increases the variance of the underlying population with the condition.

To try and improve the accuracy of our classifier, we implemented Adaboost, a boosting algorithm common in Machine Learning that combines the pooled results of many 'weak' classifiers (>50%) in order to come up with a 'stronger' classifier with higher accuracy. The end result of our efforts was 75% precision and recall. While we are pleased with this result, it is important to note that we did not come in with any machine learning expertise and built the platform in such a way that someone with more experience could plug in a classifier that would be better equipped to solve the stated problem. A summary slide on our Machine Learning approach can be found below:



Clinician-Facing User Interface:

The final component of our product was the interface with which a clinician could get the resulting data from the signal processor and classifier. Our proof of concept is shown below:



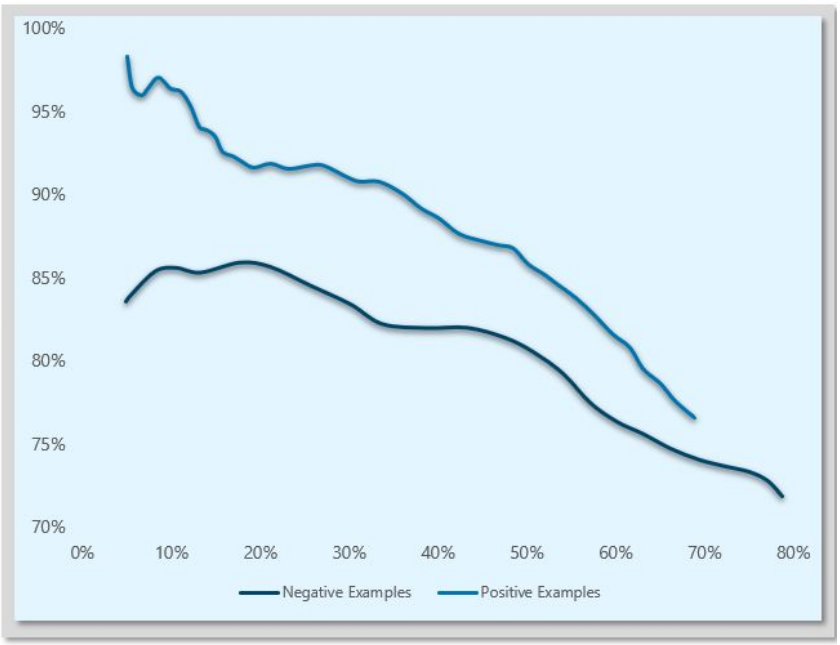
On the input screen, the doctor simply inputs the patient's name and ID and uploads the three data files associated with his sleep study. The calculation is then performed on the backend and the doctor is brought to the results screen on the right. On the top, they can see the prediction for various conditions, this is the main output of the signal processor and classifier. On the bottom left, they can use a built-in EDF viewer to scroll through the actual patient data. The tabs on the right of this view are shortcuts to areas of interest that the signal processor and classifier identified in the data. On the right the clinician has

the opportunity to create an output document which generates a report in line with those sent back from sleep clinics and containing the relevant information stored in EPIC.
Overall, we wanted to keep our UI simple and intuitive, and while we do not believe this to be a finished product, it effectively serves its purpose as a proof of concept a doctor could work with to use our platform.

6. Evaluation

We divided our evaluation phase into statistical and clinical evaluation.

Statistical:
As mentioned before, our classifier predicts COPD with a 75% precision and recall, with a graph below of the recall trade-off forfeited to increase precision:



While there isn't necessarily a 'clinically required' precision for FDA approval, we can achieve 90% precision if we drop recall to 30% - essentially only predicting yes for those with a high probability of COPD. As can be seen in the Business Analysis, 75% precision and recall results in massive savings already for the US Health System, and deploying with this accuracy already has tremendous potential to save money and lives. We will be handing off this project to our faculty advisor, who's research team has the expertise and new methods required to further increase our accuracy and help save more lives. Additionally, our clinical advisor is willing to demonstrate the efficacy of our platform by integrating it with the 5k sleep studies run at Penn every year, which will give us a large enough dataset to pursue FDA approval.

Medical:
We developed our platform with the clinician's workstream in mind. From initial conversations, we knew that the most convenient solution would integrate with the existing software platforms used to view the data such as Sandman. However, we could not find any documentation or API online on how to work

with Sandman, and therefore decided to develop a proof of concept separate application that doctors could boot up on their own system and run.
The initial reaction of doctors to our platform was very good. Doctors were happy that we worked with EDF files which made the platform flexible to integrate across different health systems. Additionally, the 'clear presentation of results' and 'intuitive and simple design' were pluses that got them excited about the platform. While they acknowledged that a plugin where they would not have to manually upload the files would be best, they were excited about the potential for our platform and were willing to work with it now in preparation for future development.

In conclusion, we are very happy with how our final project turned out. We would like to thank the CIS 350 team as well as our Faculty, Clinical and M&T Advisor for their help for all the help with our project, and hope Sleep-Dx can enable/inspire solutions that save American lives.

7. Individual Contributions

Samuel Lerner - Built the classifier (selected signals, designed architecture, implemented boosting etc…) and report-generating functionality for the platform
Kanishka Rao - Created presentations & reports, developed business analysis, lead evaluation.
Zane Stiles - Led the data cleaning signal processing efforts in MATLAB, integrated GUI with MATLAB signal processing and python classifier.
Mike Terracciano - Worked with data cleaning and focused on the platforms modular code base as well as GUI.

8, Business Analysis

Need & Value Proposition:
Chronic Obstructive Pulmonary Disease (COPD) is the 3rd largest cause of death in America. It affects an estimated 14% of adults in the United States between the ages of 40 and 79 years, but only 12 million are diagnosed. The treatment cost of COPD is cited at around $75B each year domestically, and direct healthcare costs for patients have increased almost 40% in the last 30 years. While COPD is so far incurable, late diagnosis is a contributing factor to exacerbations of COPD conditions and an increasing number of fatalities every year.
12 million adults in the U.S. are thought to have undiagnosed COPD. Additionally, Dr. Gooneratne and sleep physicians at the Penn Sleep Center believe all relevant data to diagnose someone with COPD is gathered in a sleep study - a relatively common and inexpensive test. Our platform can provide early COPD diagnosis to the 12 million undiagnosed Americans any time they subject to a sleep study for related or unrelated symptoms.
Earlier diagnosis and treatment not only improves patient outcomes and quality of life, but can also lessen the financial burden associated with the disease by a factor of 10.

Stakeholders:
The Healthcare system is a complicated relationship of various players that are all relevant to our platform:

- Patients desire the best care available and want the improved outcomes & quality of life associated with early diagnosis. They also want to avoid financial burdens associated with acute exacerbations and misdiagnosis.
- Doctors want their patients to be healthy, but also want to produce revenue. If they are able to diagnose more diseases with a sleep study, they are able to charge patients more per study, as well as prescribe more sleep studies overall to patients with a variety of symptoms.
- Health Systems want to avoid misdiagnosis of patients that waste resources, and especially in pre-operative surgical procedures, want to avoid misdiagnoses that lead to expensive malpractice lawsuits. At Penn there was recently an orthopaedic lawsuit for a patient misdiagnosed with COPD who passed away from surgery complications.
- Insurance Companies want to avoid worst-case scenarios and reduce the total direct healthcare cost for their patients (PMPM costs). They typically cover preventive & diagnostic procedures/tests.
- Software Companies such as Sandman want more people integrated with their platform. If our platform can be built as a feature/plugin for them, they may be willing to pay a royalty and charge customers for the added functionality.

Market Size:
There are roughly 3.5M sleep studies conducted each year in the United States. As stated before, there are an expected 24 million Americans with COPD, with 12 million of them currently undiagnosed. In the Swedish ARCTIC study, it was estimated that late diagnosis resulted in an average cost increase of 3,000 EUR/yr. Using the base rate of 3.7% of Americans being diagnosed with COPD, we find that we have the potential to diagnose ~97k new patients each year with 75% precision and recall. . Using the measure of $3700 USD * 97k, our platform has to potential to generate **>$350M** in savings this year for the US Health System.
The number of COPD patients increases each year as people living longer has modelled the same trend.

Competition:
There currently are not any similar Machine Learning approaches to diagnosing non-sleep-related conditions from a sleep study. The current method of diagnosis for COPD is a spirometry breathing test (pulmonary function). Currently, these range from $50-$400.
Clinically, we use more factors including medical history in our diagnosis, as well as all the data taken in by a Spirometer. To demonstrate economic superiority to the competition, we will have to price lower and demonstrate value within one year.

Path to Market:
The FDA recently launched a Digital Health Innovation Action Plan which relaxes the requirements for low-risk tech solutions to be FDA approved. Since Sleep-Dx and the resulting spirometry test have no adverse effects to a patient, the product is low-risk and eligible for this fast-tracked approval. Once approved we plan on tackling the segmented market by selling both to the sleep clinics (of which there are over 1600 nationwide) and also the 9 large software providers that are used to administer and view sleep studies. Integrating with these software products gives us the opportunity to expand our reach to more patients in this fragmented sleep clinic market.

Revenue Model/Costs

We plan to model our service as a SaaS service. On our platform, a user can upload their EDF (sleep-study data) and submit it. The result will a be a prediction of COPD for the physician to use. We have flexibility on what we can charge, but will use $75 per use for this analysis (~5% of total sleep study cost). Our costs are negligible (less than $1 per run), as running the classifier on a single patient is nearly instantaneous, and all computation and storage will happen on the cloud. The SaaS model will allow us to scale rapidly with minimal staffing, and integrating into any of the dozens of sleep studies performed each week will allow for quick deployment. We will need to demonstrate clinical efficacy by integrating into clinical trials in order to convince Insurance companies to cover the process.

However, after demonstrating clinical efficacy through clinical trials, we believe we can grow our revenue quickly.

Sleep-Dx

Revenue Projections

	Year 1	Year 2	Year 3	Year 4	Year 5
Major hospitals	1	3	3	7	8
studies per year per hospital	5,000	5,000	5,000	5,000	5,000
studies from major hospitals	5,000	15,000	15,000	35,000	40,000
Minor hospitals	0	10	10	15	20
studies per year per hospital	1,000	1,000	1,000	1,000	500
studies from minor hospitals	0	10,000	10,000	15,000	10,000
Software companies	0	0	1	2	4
studies per year per company	50,000	50,000	50,000	50,000	75,000
studies from software companies	0	0	50,000	100,000	300,000
Total Units	5,000	25,000	75,000	150,000	350,000
Price per Unit	$75.00	$75.00	$75.00	$75.00	$75.00
Gross Revenue	$375,000	$1,875,000	$5,625,000	$11,250,000	$26,250,000
Annual Savings to Insurers	$360,000,000	$380,160,000	$401,448,960	$423,930,102	$447,670,187
CAGR	5.6% [1]				
TAM Share	0.1%	0.5%	1.4%	2.7%	5.9%

1 https://www.advisory.com/daily-briefing/2017/02/16/spending-growth

We broke out our revenue over three customer segments, major hospitals, minor hospitals, and software companies. We believe we can gain one major hospital (Penn) as a customer in the first year, prove ourselves with a few other major hospitals and many minor ones, then sell into the software companies. Finally, while revenue of $26 million in year 5 sounds high, we believe it is reasonable as it is only 6% of TAM.

Finally, we believe this platform is simply proving the concept of ML diagnosis on sleep data, and we can expand to a variety of other conditions easily to rapidly scale revenue.

9. Sources

Blanchette CM, Gross NJ, Altman P. Rising Costs of COPD and the Potential for Maintenance Therapy to Slow the Trend. *American Health & Drug Benefits*. 2014;7(2):98-106.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4049119/
National Institute of Health
https://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=77
Hospital for Veteran's Administration
https://www.va.gov/health/aboutVHA.asp

Samuel Kuna 2015 VA Presentation
Kjell Åke Larsson, Christer Janson, Björn Ställberg, Karin Lisspers, Petter Olsson, KonstantinosKostikas, Milica Uhde, Leif Jorgensen, Patrik Sobocki, Gunnar Johansson Extent and impact of late vs early stage COPD diagnosis in the Swedish ARCTIC study. *European Respiratory Journal.* Sep 2016, 48 (suppl 60) PA3933;
http://erj.ersjournals.com/content/48/suppl_60/PA3933
Raluy-Callado, M., Lambrelli, D., MacLachlan, S., & Khalid, J. M. (2015). Epidemiology, severity, and treatment of chronic obstructive pulmonary disease in the United Kingdom by GOLD 2013. *International Journal of Chronic Obstructive Pulmonary Disease*, *10*, 925–937.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4437524/
Guarascio, A. J., Ray, S. M., Finch, C. K., & Self, T. H. (2013). The clinical and economic burden of chronic obstructive pulmonary disease in the USA. *ClinicoEconomics and Outcomes Research: CEOR*, *5*, 235–245.
http://doi.org/10.2147/CEOR.S34321
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3694800/
World Health Organization
http://www.who.int/respiratory/copd/burden/en/
Center for Disease Control
https://www.cdc.gov/healthcommunication/toolstemplates/entertainmented/tips/ChronicRespiratoryDisease.html
NovaSom
https://www.novasom.com/why-a-home-sleep-test/economics-of-home-sleep-testing/
Centers for Disease Control and Prevention
https://www.cdc.gov/copd/index.html
Cost Helper Health (Spirometry costs)
http://health.costhelper.com/pulmonary-function-tests.html
FDA Digital Health Innovation Action Plan
https://www.fda.gov/downloads/medicaldevices/digitalhealth/ucm568735.pdf
Sleep Review Magazine
http://www.sleepreviewmag.com/2016/06/sleep-services-market-grow-cagr-12-9-2021/