

CS Bits & Bytes is a bi-weekly newsletter highlighting innovative computer science research. It is our hope that you will use CS Bits & Bytes to engage in the multi-faceted world of computer science to become not just a user, but a creator of technology. Please visit our website at: <http://www.nsf.gov/cise/csbytes>.

September 24, 2012

Volume 2, Issue 2

Machine Learning Saves Babies!

Computing algorithms save the lives of premature babies! More than 500,000 babies are born prematurely (before 37 weeks) in the United States every year. These babies have an increased risk of major health complications, including death within their first year of life.

MUST SEE!



Learn more about Dr. Saria's Research and the PhysiScore at: <http://engineering.jhu.edu/new/images2/MUCMD2011-SUCHI-SARIA.mp4> Video courtesy Clarence Wigfall/Advancedsciencecommunications.com. Video courtesy Clarence Wigfall/Advancedsciencecommunications.com.

In most hospitals, babies born prematurely are immediately brought to the Neonatal Intensive Care Unit (NICU) for **evaluation and monitoring** (and treatment/interventions as needed) due to their increased risks. Once in the NICU, these babies undergo a multitude of tests to keep track of the health of the baby, including continuous monitoring using **sensors** attached to their body that measure physiological status (heart rate, respiration rate, oxygen saturation rate, blood pressure, etc.). For years, this data was quickly scanned by doctors, courses of treatment were determined, and then the data was never used again.

Computer scientists are now using machine learning tools to look at the massive amounts of data to find common patterns among babies, creating new measures for predicting decline in the babies' health,

and forming treatment plans. Machine learning is a discipline of computer science devoted to the development of algorithms that take raw data as input and find patterns or make predictions based on features of the underlying data. Machine learning has many applications, including finding volcanoes on the surface of Mars and determining defects on the surface of semiconductor chips.

The **PhysiScore** uses non-invasively measured data from the first few hours of life and produces a probability score for each baby that represents the babies overall illness severity and likelihood of developing major complications downstream. This score is like the APGAR (Appearance, Pulse, Grimace, Activity, and Respiration) score administered minutes after a baby is born in that it helps doctors and nurses evaluate the health of newborn babies. However, the PhysiScore is significantly more accurate, and since it can be easily automated on existing monitoring devices, it makes the incorporation of such a score easy within the clinical workflow. It is also more accurate without being invasive (no need for blood draws, spinal taps, etc.). The doctors can use this information to plan for infant transport and plan patient management, thereby providing them with greater chances for future healthy living.

As health systems data are increasingly digitized and methods allow for a better sense of the multitude of noisy data, every aspect of care for



Dr. Saria and the team analyzing data in the NICU. Photo from: <http://www.eurekalert.org/multimedia/pub/25303.php?from=167838>

individuals that now need extensive support from the health system stands to benefit from these approaches.

Who Thinks of this Stuff?! Suchi Saria recently joined the computer science and school of public health faculty at Johns Hopkins University, where she hopes to innovate computational systems that utilize observational data to help inform and improve the delivery of care. In her spare time, Dr. Saria enjoys reading and spending time outdoors biking, hiking, or camping. She's also just starting to learn how to fly a helicopter!

Links:

Learn about Machine Learning at: <http://robotics.stanford.edu/~nilsson/mlbook.html>.

Read the full scientific paper about the PhysiScore at: <http://stm.sciencemag.org/content/2/48/48ra65.abstract>.



Image of Dr. Suchi Saria.

Activity:

This activity was adapted from Su, Francis E., et al. "Medical Tests and Bayes' Theorem." Retrieved from: <http://www.math.hmc.edu/funfacts>.

Read the following to your class:

Suppose you are worried that you have a virus. You decide to get tested. Now suppose the testing methods are accurate 99% of the time (regardless of whether the results come back positive or negative). Suppose the virus is actually a rare condition, occurring in only 1 of every 10,000 people (.01%).

If your test results come back positive, what are the chances you really do have the disease?

(a)99% (b)90% (c)10% (d)1%

Are you surprised that the correct answer is (d)1%?

In small groups, have students discuss the following questions:

1. Why aren't test results reliable?
2. Would the result be so surprising if the virus was more common?
3. How would your chances change if the percentage of false positives and false negatives were different?
4. Does the presence (or absence) of symptoms have any bearing on the result?

Further exploration: As a class, construct a tree diagram or table representing the possible probabilities. Then use the representation to compute the answer above.

See <http://www.math.hmc.edu/funfacts/ffiles/30002.6.shtml> for an explanation of how Bayes' Theorem provides a method for computing the probability of event A, given that event B has happened.

To see an additional demonstration of Bayes' Theorem, look at the Interactive Probability Computation of Being Sick After Having Tested Positive for a Disease at: <http://demonstrations.wolfram.com/ProbabilityOfBeingSickAfterHavingTestedPositiveForADiseaseBa/>

CS BITS & BYTES

<http://www.nsf.gov/cise/csbytes/>

Please direct all inquiries to: csbitsandbytes@nsf.gov

National Science Foundation
Computer & Information Science & Engineering Directorate
4201 Wilson Blvd Suite 1105
Arlington VA, 22230