Analysis of acoustic features for speech sound based classification of asthmatic and healthy subjects

Shivani Yadav[1], Merugu Keerthana[2], Dipanjan Gope[3], Uma Maheswari K.[4], Prasanta Kumar Ghosh[5]

[1] BioSystems Science and Engineering, Indian Institute of Science (IISc), Bangalore, India
[2] Rajiv Gandhi University of Knowledge Technologies, Kadapa, Andhra Pradesh, India
[3] Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India
[4] Pulmonary Medicine, St. Johns National Academy of Health Sciences, Bangalore, India
[5] Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India

Poster Sesssion: TU2.PB.12: Speech Analysis and Coding
Tuesday, 5 May, 16:30 - 18:30

Table of Contents

Introduction

Objective

- Performance analysis of speech sounds for asthamatic and healthy subject classification by **INTERSPEECH 2013 Computational Paralinguistics Challenge baseline (ISCB)** features.

Introduction

Objective

- Performance analysis of speech sounds for asthamatic and healthy subject classification by **INTERSPEECH 2013 Computational Paralinguistics Challenge baseline (ISCB)** features.

Why ISCB features?

Introduction

Objective

- Performance analysis of speech sounds for asthamatic and healthy subject classification by **INTERSPEECH 2013 Computational Paralinguistics Challenge baseline (ISCB)** features.

Why ISCB features?

1. *Iwarsson et al. [1], Dogan et al. [2], Hamdan et al. [3] has showed acoustic characteristics of speech sounds differ between asthmatic and healthy subjects.*

Introduction

Objective

- Performance analysis of speech sounds for asthamatic and healthy subject classification by **INTERSPEECH 2013 Computational Paralinguistics Challenge baseline (ISCB)** features.

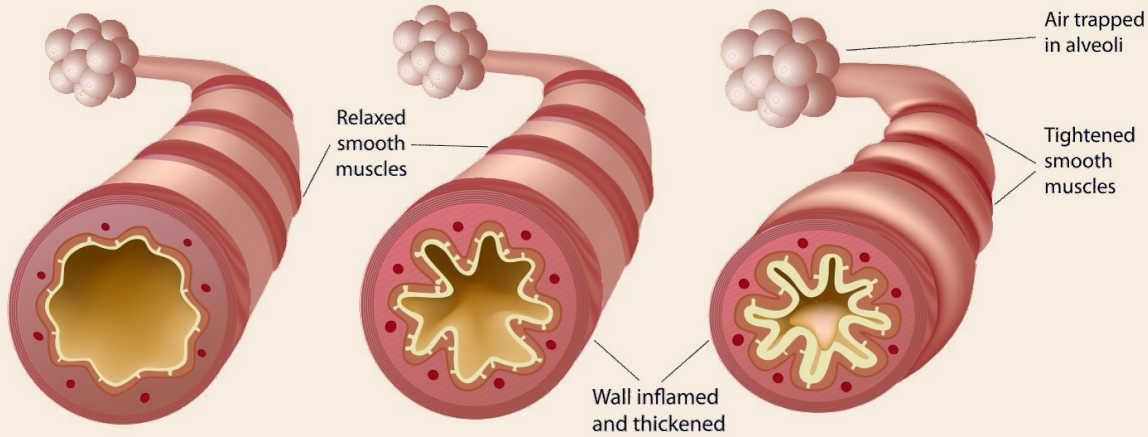
Why ISCB features?

1. *Iwarsson et al. [1], Dogan et al. [2], Hamdan et al. [3] has showed acoustic characteristics of speech sounds differ between asthmatic and healthy subjects.*

Spectral, cepstral, energy, *and* **excitation/source** *information all are present in ISCB feature set[4].*

What is Asthma??

1. Asthma is an **inflammatory disease** of the airways resulting in a number of symptoms including **obstruction of the airways**, chest discomfort or pain, cough, and wheezes or other peculiar sounds during breathing.
2. 334 million is global burden of asthma[2].



Healthy and Asthmatic Airway [1]

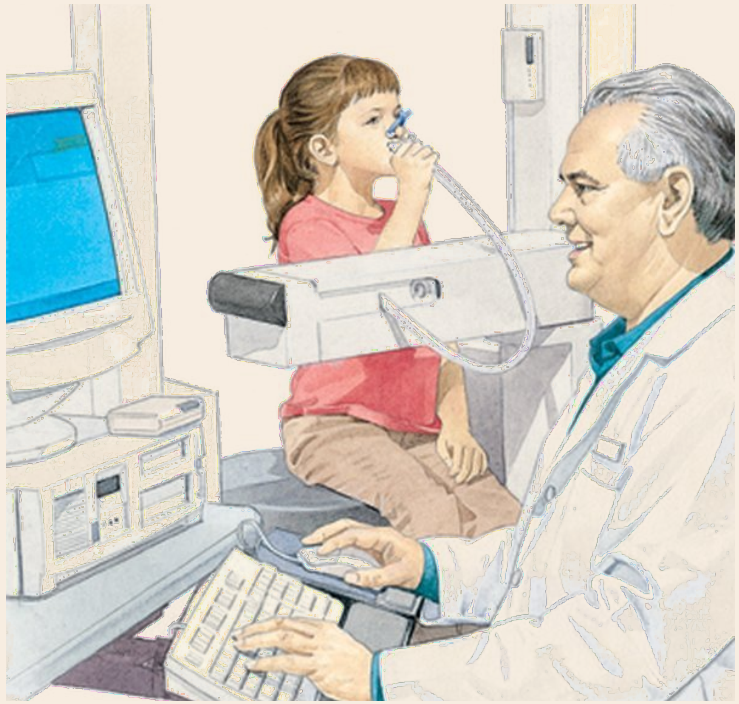
[1] http://ib.bioninja.com.au/options/option-d-human-physiology/d6-transport-of-respiratory/asthma.html

[2] http://www.globalasthmareport.org/burden/burden.php

Table of Contents

Motivation

- **Spirometry** is a lung function test that **measures how much and how fast a person can exhale air**.
- Helps in **diagnosis** and **monitoring** asthma in hospital.
- Drawbacks of spirometry
  1. Very strenous
  2. More training required
  3. More time required to screen large population



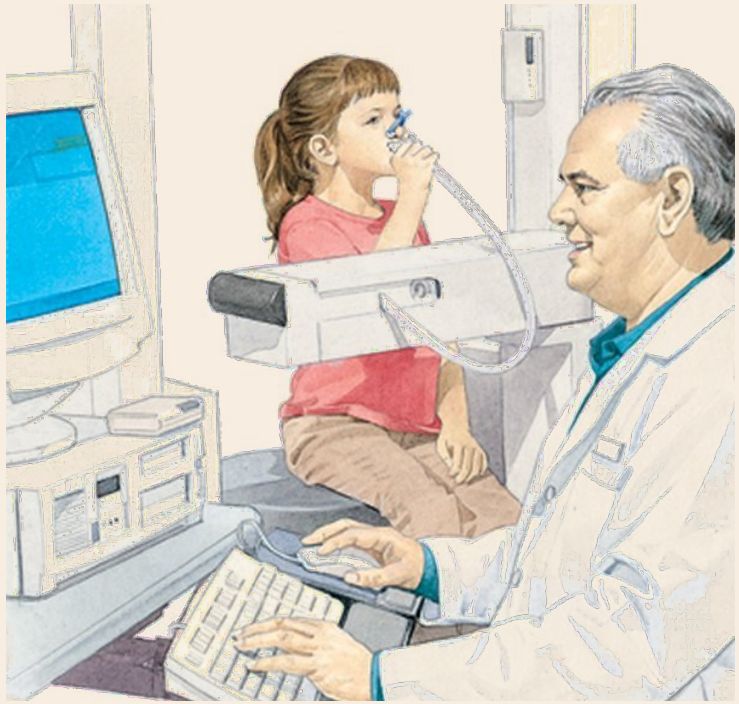
Spirometry [1]

[1] http://www.virtualimaging.org/pft.html

Motivation

- **Spirometry** is a lung function test that **measures how much and how fast a person can exhale air**.
- Helps in **diagnosis** and **monitoring** asthma in hospital.
- Drawbacks of spirometry
  1. Very strenous
  2. More training required
  3. More time required to screen large population

Is there a simpler approach for asthma diagnosis?

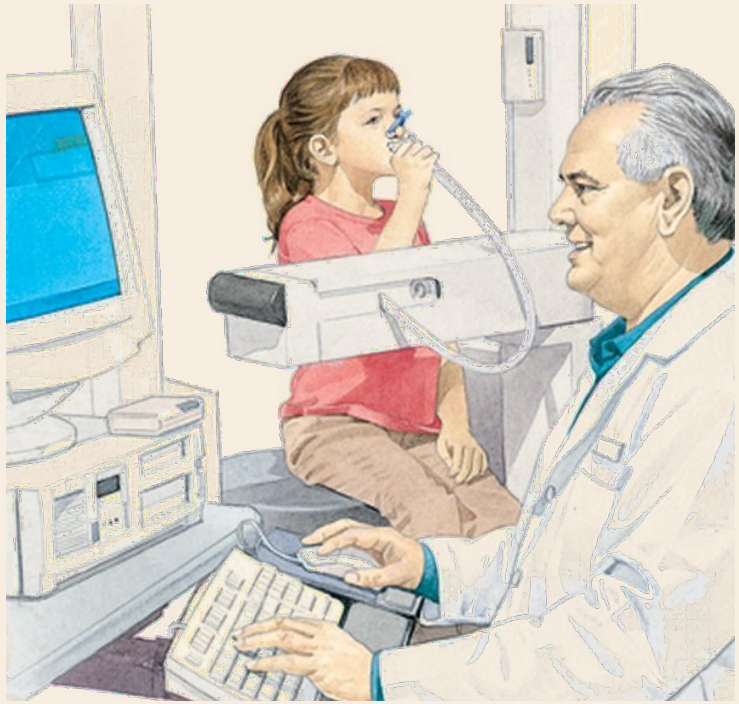


Spirometry [1]

[1] http://www.virtualimaging.org/pft.html

Motivation

- **Spirometry** is a lung function test that **measures how much and how fast a person can exhale air**.
- Helps in **diagnosis** and **monitoring** asthma in hospital.
- Drawbacks of spirometry
  1. Very strenous
  2. More training required
  3. More time required to screen large population



Spirometry [1]

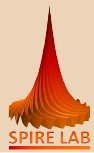
Is there a simpler approach for asthma diagnosis?

Yes, **Sound based analysis** *can be.*

[1] http://www.virtualimaging.org/pft.html

Table of Contents

Dataset Description



- 95 subjects
  - Patients: 47 (28M, 19F)
  - Control: 48 (24M, 24F)
- Age range: Patients 15-71 years, Controls 19-60 years.
- Stimulus used
  - Speech sounds: /ɑː/ (as in 'After'), /iː/ (as in 'Eat'), /uː/ (as in 'Cute'), /eɪ/ (as in 'Pay'), /oʊ/ (as Only'), /s/ (as in 'Same'), and /z/ (as in 'Zoom')
  - Non speech sounds: cough, wheeze and inhale and exhale.
- Each stimulus was recorded on an average five times per subject.
- Sampling rate: 44.1 kHz.

Table of Contents

Experiments & Questions addressed

Exp1. *Comparsion of ISCB and baseline Mel-frequency cepstral coefficients (MFCC) features performance for Asthmatic and Healthy subjects classification task.*

1. **How** does **ISCB** features perform **compared** to the baseline **MFCC** features classification?
2. **Which stimuli** (cough, wheeze and speech sounds) is better for classification?

Exp2. *Forward feature group selection for speech sounds.*

1. **Which feature group** in ISCB feature sets selected after forward feature group selection in each speech sound?
2. Is there any **feature group common** across all the speech sounds after forward feature group selection?

Exp3. *Forward feature selection in best selected group of Exp2 in all speech sounds.*

1. **Which features** are selected after forward feature selection in best feature group?

Experimental setup

- ISCB features computation: analysis window: 20ms, shift : 10ms except for F_0 which uses 60ms window with 10ms shift.
- ISCB feature groups into 2 sets: SetA has 5900 features, SetB has 473 features and total 6373 features.Total 21 groups.

Table: List of ISCB groups (number of features) .

	SetA	SetB
GROUPS	Loudness (100), Modulated loudness(100), Root mean square (RMS) Energy(100), Zero crossing rate (ZCR) (100), RASTA auditory bands (2600), MFCC (1400), Band energy (200), Spectral Roll Off (400), Spectral flux (100), Spectral centroid (100), Spectral entropy (100), Spectral moments (300), Spectral slope (100), Harmonicity (100), Spectral Sharpness (100)	Fundamental frequency(F_0) (83), Probability of voicing (78), Jitter (78), Jitter of Jitter (JJ) (78) Shimmer (78), logarithmic harmonic to noise ratio (LHNR) (78)

Experimental setup

- Support vector machine (SVM) with RBF kernel is used in 5 fold cross-validation setup.
- **Baseline**: MFCC statistics features with SVM as classifier[1].
- Evaluation metric
  - Total classification accuracy

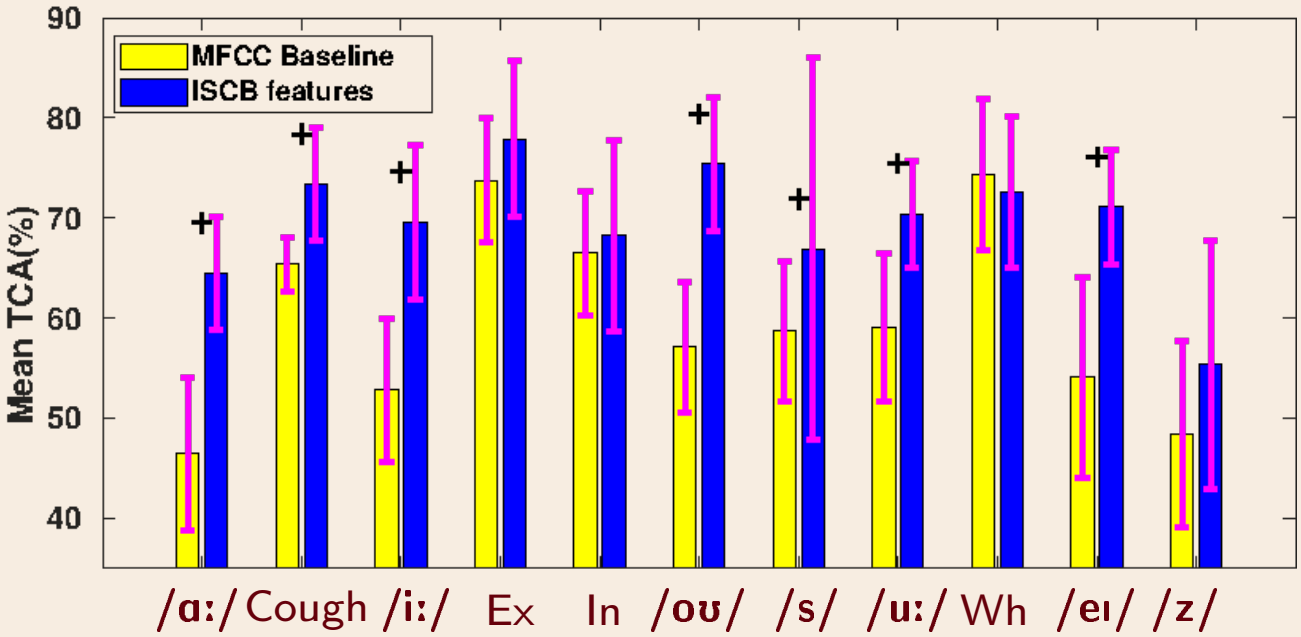
$$TCA = \frac{TP + TN}{TP + TN + FP + FN}$$

where, TP (True Positive), TN (True Negatives), FP (False Positives) and FN (False Negatives).

[1] Shivani Yadav, NK Kausthubha, Dipanjan Gope, Uma Maheswari Krishnaswamy, and Prasanta Kumar Ghosh, "Comparison of cough, wheeze and sustained phonations for automatic classification between healthy subjects and asthmatic patients," in 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 1400–1403.

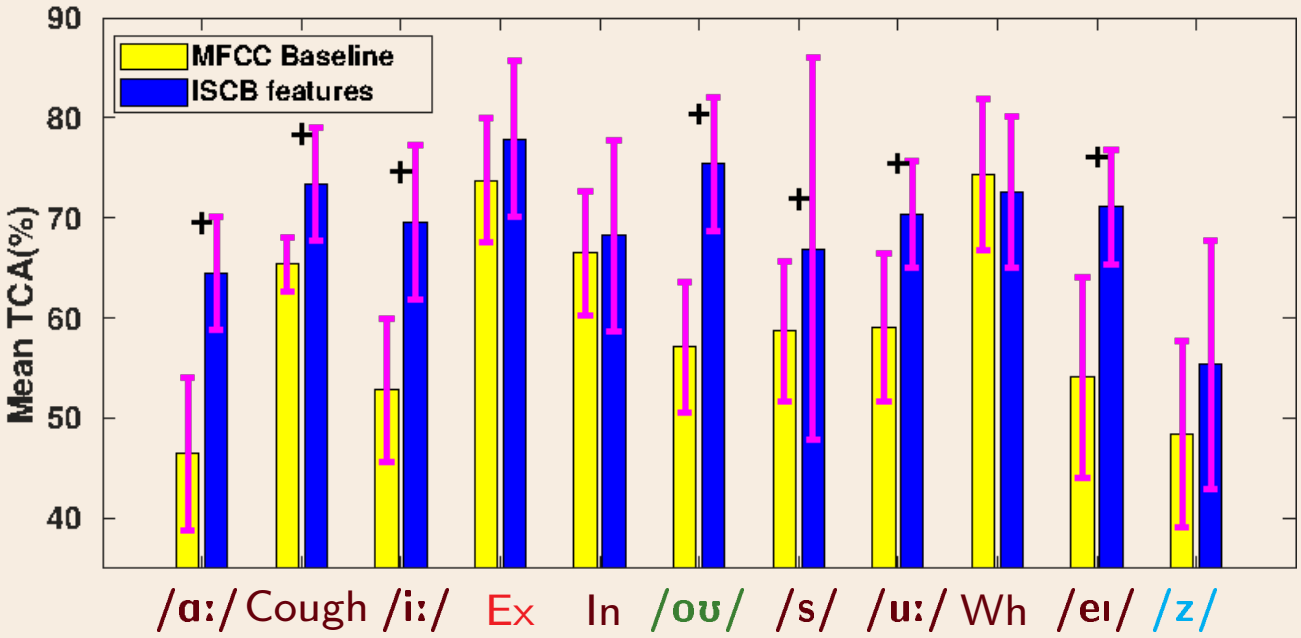
Table of Contents

Exp1. Comparison of ISCB and baseline features



1. How does ISCB features performs compared to the baseline MFCC features for the asthma vs healthy classification task?
   - **ISCB** features **performed better** than the baseline in all stimuli except wheeze.
   - All **vowels** sounds showed an **significant improvement** in TCA of atleast 10% and by fricatives atleast 6%.

Exp1. Comparison of ISCB and baseline features



2. Which stimuli among speech and non-speech sounds performed best for the classification?
   - **Exhale** performed the best among all sounds.
   - Among speech sounds /oʊ/ peformed best which is comparable to Exhale.
   - /z/ performed poor among all speech and non-speech sounds.

Exp2. Forward feature group selection

Table: Best selected feature groups for speech stimuli.

Stimuli	Best selected feature groups
/ɑː/	MFCC
/iː/	Jitter of the Jitter, MFCC
/oʊ/	Loudness, MFCC
/s/	Loudness, Spectral Entropy, Spectral Flux
/uː/	MFCC
/eɪ/	MFCC
/z/	Logarithmic Harmonic to Noise ratio

3. Which feature group in ISCB feature sets selected after forward feature group selection in each speech sounds?
   - **Stimuli dependent.**
   - **Loudness** and **MFCC** are selected for the best performing stimuli /oʊ/.
4. Is there any feature group common across all the speech sounds after forward feature group selection?
   - **MFCC** is common across all the **vowels** stimuli.

Exp3. Forward feature selection in best group

Table: Best selected feature in best performing groups for speech stimuli.

Stimuli	Best selected feature
/ɑː/	Interquartile range between 2^{nd} and 3^{rd} quartile of MFCCs 2^{nd} coefficient
/iː/	1^{st} quartile of MFCCs 2^{nd} coefficient
/oʊ/	Interquartile range between 2^{nd} and 3^{rd} quartile of mel-scale loudness
/s/	linear regression 2nd coefficients of mel-scale loudness
/uː/	linear prediction gain of MFCCs delta 8^{th} coefficient.
/eɪ/	1^{st} percentile of MFCCs 3^{rd} coefficient
/z/	Skewness of derivative of logarithmic harmonic to noise ratio

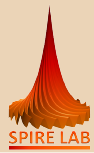
5. Which features are selected after forward feature selection in best feature group?
   - **MFCC group** features are the best performing in /ɑː/, /iː/, /uː/ and /eɪ/ speech sounds.
   - For best performing **/oʊ/** sound, **mel-scale loudness group** feature is selected at the top.

Table of Contents

Conclusion and Future work

1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.

Conclusion and Future work



1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.
2. **Exhale** performed best among all stimuli for asthma and healthy classification and /**ʊ**/ performed best among all speech sounds.

Conclusion and Future work

1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.
2. **Exhale** performed best among all stimuli for asthma and healthy classification and **/oʊ/** performed best among all speech sounds.
3. Feature group selected across speech sounds is **stimuli dependent**, although **vowels** have **MFCC** as common group.

Conclusion and Future work

1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.
2. **Exhale** performed best among all stimuli for asthma and healthy classification and **/oʊ/** performed best among all speech sounds.
3. Feature group selected across speech sounds is **stimuli dependent**, although **vowels** have **MFCC** as common group.
4. **MFCC** and **loudness** is selected as best performing group for **/oʊ/** stimuli.

Conclusion and Future work

1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.
2. **Exhale** performed best among all stimuli for asthma and healthy classification and **/oʊ/** performed best among all speech sounds.
3. Feature group selected across speech sounds is **stimuli dependent**, although **vowels** have **MFCC** as common group.
4. **MFCC** and **loudness** is selected as best performing group for **/oʊ/** stimuli.
5. **MFCC group features** are best selected among all **vowels** sounds except **/oʊ/** which has **Interquartile range between 2^{nd} and 3^{rd} quartile of mel-scale loudness** as the best feature.

Conclusion and Future work

1. **ISCB** features showed **improvement over baseline** features in all sounds except wheeze.
2. **Exhale** performed best among all stimuli for asthma and healthy classification and **/oʊ/** performed best among all speech sounds.
3. Feature group selected across speech sounds is **stimuli dependent**, although **vowels** have **MFCC** as common group.
4. **MFCC** and **loudness** is selected as best performing group for **/oʊ/** stimuli.
5. **MFCC group features** are best selected among all **vowels** sounds except **/oʊ/** which has **Interquartile range between 2^{nd} and 3^{rd} quartile of mel-scale loudness** as the best feature.
6. We want to further explore the use of ISCB features for **running speech.**

Table of Contents

References

1. Jenny Iwarsson, Monica Thomasson, and Johan Sundberg, "Effects of lung volume on the glottal voice source," Journal of voice, vol. 12, no. 4, pp. 424–433, 1998.
2. Muzeyyen Dogan, Emel Eryuksel, Ismail Kocak, Turgay Celikel, and Mehmet Ali Sehitoglu, "Subjective and objective evaluation of voice quality in patients with asthma," Journal of Voice, vol. 21, no. 2, pp. 224–230, 2007.
3. Abdul Latif Hamdan, Georges Ziade, Maher Kasti, Leslie Akl, Ibrahim Bawab, and Nadim Kanj, "Phonatory symptoms and acoustic findings in patients with asthma: a cross-sectional controlled study," Indian Journal of Otolaryngology and Head and Neck Surgery, vol. 69, no. 1, pp. 42–46, 2017.
4. Florian Eyben, Real-time speech and music classification by large audio feature space extraction, Springer, 2015.

Table of Contents

Authors thank the **Achuth Rao MV** for his valuable inputs and the **Department of Biotechnology, Govt. of India** for their support in this work.

THANK YOU

Questions

Have Questions/suggestions? Write to us at **spirelab.ee@iisc.ac.in**