Mphasis
The Next Applied

# Deep Learning and Image Analytics to Monitor Healthcare Guidelines and Workplace Safety

Whitepaper by Manish Shukla, Assistant Manager – Data Science |
Narayan Mishra, Assistant Manager – Data Science |
Vibha Bhagchandani, Associate Vice President, AI & Cognitive Solutions, Mphasis NEXT Labs

# Contents

# 1.
# Introduction

COVID-19 has impacted lives and business operations across the globe severely. It can be transmitted from one person to another based upon the proximity. To avoid the spread and transmission of COVID-19, guidelines published by WHO mandate wearing a mask and maintaining social distancing of at least 1 meter in public spaces. To overcome the challenges posed by COVID-19 pandemic, we must change the way we work. This will ensure the safety and wellbeing of the employees and thus help in maintaining business continuity. There are preventive measures stated by WHO and Governments for offices that should be strictly followed to contain the spread of COVID-19. Some of the preventive measures are that individuals must maintain a minimum distance of 1 meter in common places as far as possible, always use face covers/masks, maintain hand hygiene, etc. By enforcing these preventive measures, spread of COVID-19 can be contained within office premises and wellbeing of employees, their families and critical business operations can be ensured.

Using Machine Learning, some of these preventive measures such as social distancing, covering face with mask can be enforced efficiently without human intervention and records can be maintained for statistical purpose and can be visualized over a dashboard for further analysis. These solutions can help organizations to manage office premises in a better way so that business continuity can be maintained, and loss of revenue can be averted. Monitoring the adherence to these guidelines while workers are coming to office can become a challenge if done manually. In this paper, we propose an approach to monitoring the same by building an AI-powered system using the advancements in Machine Learning and computer vision techniques. We will explain the solution approach and how one can go about building a system that can assist in day-to-day operations.

# 2.
# Approach

Computer vision techniques have made tremendous progress through real-time processing of images and videos. Deep Learning methods require high computation, but the availability of GPUs, TPUs and the maturity and flexibility provided through various cloud service providers has made this task easier. Computer vision is the subfield of Artificial Intelligence that enables computers to mimic human visual system. Using computer vision technology, machines can derive useful information from digital images and videos. The typical steps involved in computer vision are acquisition of images or videos, converting them into computer understandable format and extracting relevant and useful information from them. Based on the extracted information, further action can be taken. Object detection, automatic driving systems and face recognition are some such application areas. These benefits of computer vision can be helpful in aiding and monitoring government guidelines to minimize the impact of COVID-19 spread. Computer vision

can be used for no touch biometric system, identifying individuals not wearing mask and for monitoring adherence to social distancing guidelines at common places within the workplace. Computer vision and Machine Learning methodologies can be leveraged to build a real-time monitoring and dashboarding system that can also store the information for compliance and audit purposes and reduce the dependency of human monitoring.

# 3.
# Introduction to Convolutional Neural Network

The Convolutional Neural Network (CNN) is one of the most often used deep neural networks. Convolution is a mathematical linear action between matrices that gives it its name. A convolutional neural network is made up of nodes that are present in input layer, one or more hidden layers and output layer. Each node is connected to the others and has a weight and threshold assigned to it. After a mathematical computation, if a node's output exceeds a certain threshold value, the node is activated, and data is sent to the next layer of the network. Otherwise, no data is sent on to the network's next layer.

Because of its improved performance, CNN may be employed with several input forms, although the most frequent applications include pictures, video, audio or voice input sources. In a CNN design, there are three types of layers: convolutional layer, pooling layer and fully connected layer. Initially, convolutional layer is present in a standard CNN, followed by further convolutional layers or pooling layers and eventually a fully connected layer. For feature engineering and intricate pattern finding, convolutional and pooling layers are employed, whereas fully connected layers are used for Machine Learning tasks like classification and regression.

The most essential component of a CNN architecture is convolutional layers. The convolution layers in the CNN are sliding filters to bring out the indirect features hidden in the input data. To capture the nonlinearity in the input data, activation functions such as Rectified Linear Unit (ReLU), Tan hyperbolic or others are used. Convolutional layer is usually followed by pooling layers which are used for reducing the dimensionality of the input vector and number of parameters.

The common pooling techniques are average and max pooling. After the input vector is processed at convolutional and pooling layers, it is propagated to fully connected layers for final processing and prediction for the task in hand. Once the prediction is made, predicted value is compared against the actual value and using backpropagation techniques, weights and biases are updated. This process is repeated several times to improve metrices like accuracy, precision, recall or squared mean error of the model.

Generally, CNNs are mostly used with image related use-cases like identifying satellite images and classifying hand-written characters. In recent times, CNNs are being used for image segmentation, object identification, face detection, face recognition, activity tracking, etc. Sometimes, CNNs are used for Natural Language Processing (NLP) related tasks also. The major advantage with a deep neural network is that a pre-trained neural network can be used by others using the concept of transfer learning.

# 4.
# Introduction to Transfer Learning

Transfer learning is the process of using knowledge gained while solving a Machine Learning or Deep Learning problem and applying it to another related problem. Transfer learning is a popular approach generally used in computer vision and natural language processing tasks to solve complex problems having less training data available. It takes benefit from pre-trained reusable models trained for a similar task and similar dataset. Transfer learning is widely used in training Deep Learning models as it helps reduce the need for very high compute and larger training datasets that are typically needed for neural networks, thereby reducing the training time as well.

If the characteristics learnt from the first tasks are generic and can be fine-tuned for other particular tasks, pre-trained models can be employed. Modeling with transfer learning entails looking for pre-trained models that can handle the problem at hand. For object identification and recognition, prominent pre-trained models include Resnet50 and Faster R-CNN; GPT and BERT for natural language processing, and FaceNet for face recognition applications. Transfer learning speeds up the training process and improves the Deep Learning model performance. The other advantage is the improved generalization of already trained models. The use of a suitable pre-trained model can eliminate the need for a large amount of data needed for training the deep neural network.

Following the approach mentioned above which includes convolutional neural networks and pre-trained models by utilizing the concepts of transfer learning, we propose a solution that can be used in office premises to classify faces with and without mask, predicting the identity of individuals not wearing face mask and giving alerts if the social distancing norms are not followed.

# 5.
# Solution Design

There are three components of the proposed solution. The first component locates the faces present in the incoming stream of input data and classifies them into 2 categories based on if the face mask is worn or not. The second component predicts the identity of individuals who are not following the face mask guidelines. The third component is used to predict if social distancing is maintained or not. Each component is described in detail in the following sections.

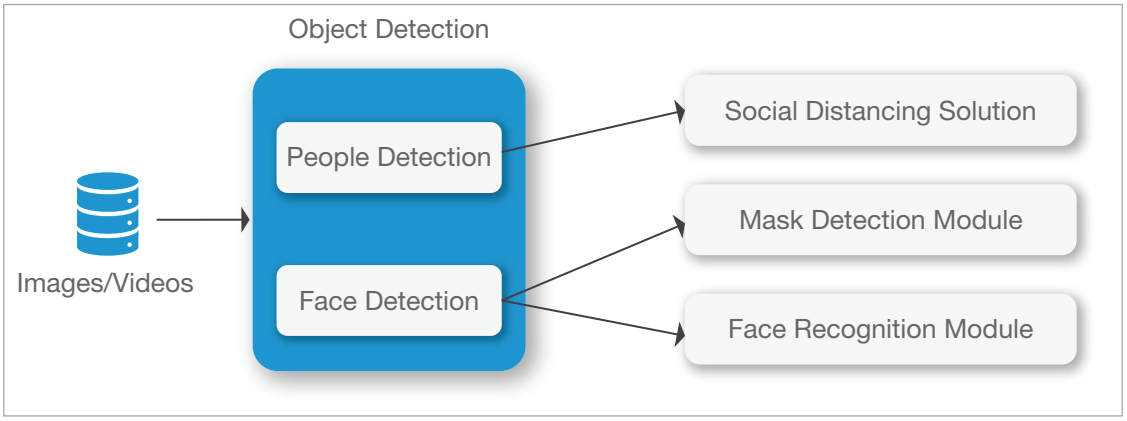


*Fig 1. Overview of Solution design*

## Mask Classifier Module

Face mask detection system can be created utilizing Deep Learning and OpenCV methods, and it can be used to monitor if individuals are wearing face masks in real time or not. The technology may be expanded to allow for real-time monitoring of numerous locations inside the office.

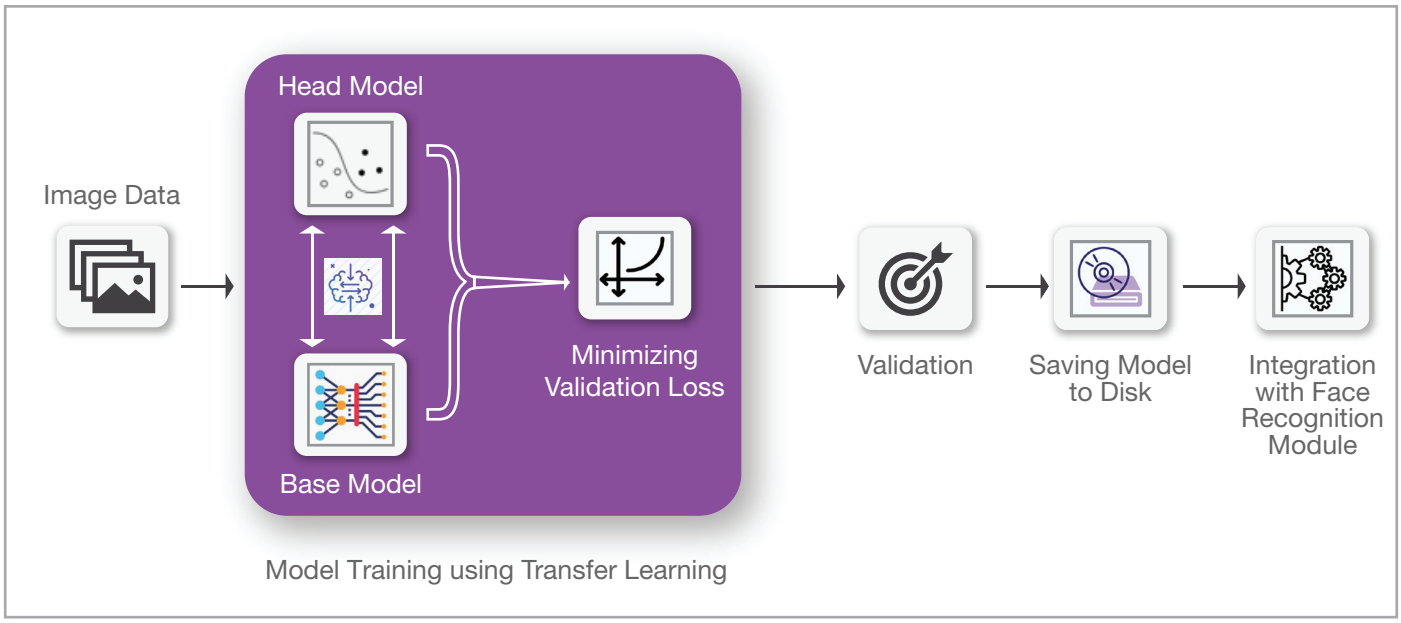


*Fig 2: Process flow diagram for face mask classification*

**The mask classifier solution can be built in the following steps:**

- The first step is to detect the faces present in the images or video streams
- The detected faces can be classified in two categories – with face mask and without face mask and accordingly images can be labelled as "Mask" or "No Mask," and then preprocessed and translated into pixel format using OpenCV
- The images in the pixel format along with the correct labels is then passed to convolutional neural network for training a supervised mask classifier model
- After the training is done, model will learn the patterns present in the data using which it can classify the face as with mask or without mask with some probability score
- The model can be saved and used in future for classifying the faces present in incoming images or video feed into “Mask” or ‘No Mask’ categories

## Face Recognition Module

Once the face is detected in incoming stream of data, face recognition solution can be used to get the identity/identities of the detected faces. This solution can be used for several purposes such as no touch attendance tracker, automatic door opening for authorized individuals and getting the identity of individuals not following mask guidelines.

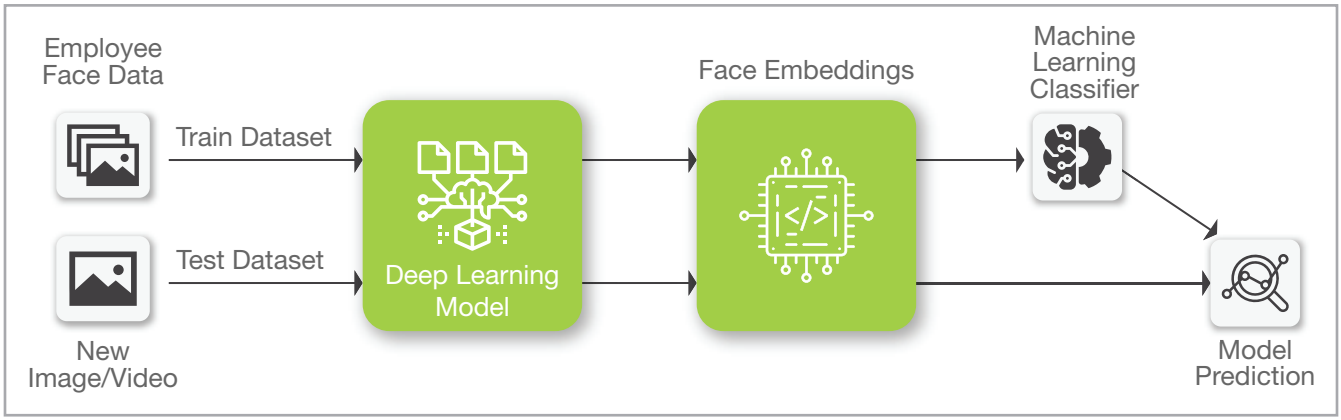


*Fig 3: Training & inference(prediction) pipeline diagram for person identification using face recognition algorithm*

**The steps involved in building a face recognition solution are:**

- Detecting faces present in an image or video feed automatically. There may be minor variances in human faces, but it is reasonable to conclude that some characteristics are shared by all human faces.
- Providing the label that could be the unique identifier for the known detected faces. This label will then be used for building a supervised face recognition model.
- The next step is to extract features present in the detected face. Face embeddings is one such method to extract facial features. Face embedding is a 128-dimensional vector obtained with the help of a neural network. A neural network takes a picture of a person's face as input and produces a vector that reflects the face's most essential characteristics. During the training process, the neural network learns to output similar vectors for faces that appear to be similar. Combining face embeddings with labels will form the training data for face recognition solution.
- Once the training data is prepared, Machine Learning or Deep Learning classification model can be trained. The model will learn the patterns present in the data and map it to the label of the individuals it belongs to.
- This trained model can be saved and used to make prediction about the identity of faces in future on new incoming images or videos

## Social Distancing Module

Distancing oneself from others to prevent the transmission of infectious diseases is known as social distancing and as mentioned in the previous sections, social distancing is critical to prevent disease spread in the times of epidemics and pandemics. It is the most effective non-pharmaceutical method of preventing the transmission of COVID-19. Social distancing can monitor using OpenCV, computer vision and Deep Learning technologies. The typical steps involved in building a social distancing solution are:

- Detect all individuals in an image or video feed using object detection
- Calculate the pairwise distances between all of the individuals found in the video/image feed
- Check to determine whether any two individuals are fewer than certain threshold pixels away based on these distances

There are pre-trained algorithms such as YOLO (You Only Look Once) that can be used for object detection task. In social distancing use case, objects are human and using YOLO they can be identified easily in a video feed.

The advantage of using social distancing detector using OpenCV and Deep Learning is that it can monitor if social distancing is maintained at multiple places in real time and can alert individuals to instill behavior change and inform concerned authorities on consistent violation, if required.

# 6.
# Application Areas and Benefits of the Solution

An AI powered facial recognition & compliance solution can be extended for further image recognition applications as in automobile security to prevent theft of important items and avoid accidents, in ride-sharing apps to guarantee the identity of driver, in smart devices to manage access-control, in immigration offices to control criminals who want to cross borders, in organizations to identify unauthorized persons tailgating or gatecrashing, in schools to track student's attendance and ensure campus safety, etc. While some industries have already adopted such solutions, many are planning to adopt AI centric face detection and recognition solutions. Some examples are retail stores, public places, public transport, airports, hotels, taxis, etc., that can implement real-time monitoring solutions using CCTV feed. Implementing the solution can impact enterprises in the following ways:

**a. Reduce manual checks and automate monitoring workflow:** Monitoring large people groups, wearing masks and maintaining social distancing is a difficult task when done manually. Leveraging new-age AI technology can help in reducing manual checks and automating monitoring of public places. It will also reduce the risk of administration teams and personnel responsible for monitoring, getting infected.

**b. Health reporting & monitoring to influence policy making:** An AI based solution can help generate real-time data that can be used to build statistical reports with various metrics like adherence to safety guidelines, number of healthy employees, number of infected employees and draw co-relations between behavior and disease spread. Such statistics collected over time not just help employers design policies that encourage and incentivize healthy behavior patterns, but can also inform policy change at large scale and help authorities prepare better for the future and reduce loss. E.g., Observing a rise of infection in employee base visiting office premises, authorities can decide to close premises before further spread, disinfect certain areas to keep things under control, provide private transport facility, facilitate and plan infrastructure to enable more employees working from home as well as provide medical support options to them.

**c. Influencing positive attitude and reducing mental health risk:** Using data available from such a system one can plan for contingencies i.e., providing employee quarantine facilities and medical assistance as well as design programs to incentivize by providing potential Wellness Bonus Leaves to employees following guidelines and not falling ill, recognize them at organization

level as Health Evangelists. This helps in not just maintaining the physical health of employees, but also keeping their morale high and mental health stable when they have an assurance that they are covered and have support in time of need. All these measures can contribute to avoiding any employee casualty, building better employee-employer relationship and build better brand reputation in the market.

# 7.
# Challenges and Considerations

There are two types of challenges in deploying such a solution at scale. The first one is establishing and maintaining data privacy and security procedures to safeguard data from theft and harmful use which is a socio-technical challenge, and the other is Accuracy of prediction and false positives which is purely a technical challenge.
The necessity of privacy and data protection is becoming increasingly acknowledged as more social and commercial activities take place online. The collection, use and disclosure of personal information to other parties without knowledge or consent is also a source of concern. According to United Nations Conference on Trade and Development, 128 of the 194 nations have enacted legislation to provide data and privacy protection. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide.

Enterprises need to establish and maintain data privacy and security procedures to safeguard data from theft and harmful use. For Machine Learning technologies, huge amount of data is required and preventing misuse of data and protecting is a major concern and topic of interest for a lot of organizations, individuals and countries. Regulations like GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act) have been instilled by countries and regions to ensure data privacy and data protection and organizations need to abide by the same. Supporting business continuity and workplace safety using Artificial Intelligence and Machine Learning requires collecting data such as facial features. This is very confidential data, and it should be protected to safeguard the employee's rights. Some of the common concerns for data safety and privacy are tampering the data, unauthorized collection and storage of data and if the regulatory obligations are followed or not. Facial features being biometric information, if not protected, can be used for multiple malicious purposes such as mass surveillance, losing privacy on global scale and security breaches. Facial information is also used as a password by other devices such as mobile phones and compromise of such information can lead to further personal and critical data loss from such devices. Another potential misuse of this data is to use data for creating deepfakes, which requires having a lot of data with various facial expressions. Deepfakes are artificial intelligence systems that exploit Deep Learning to replace the likeness of one person in video and other digital media with that of another. The name "Deepfake" is derived from the underlying technology "Deep Learning," which is a type of artificial intelligence that can train itself how to solve problems when given huge amounts of data. So, all these issues relating to data privacy, data collection, data

handling that can lead to theft of Personal Identifiable Information (PII), cybersecurity concerns are major challenges in deploying such solutions at scale.

Accuracy of prediction and false positives is another problem area. Let us consider a scenario where Machine Learning model makes prediction as “No Mask” even if the employee is wearing a mask and triggers a warning and counts an offence. Another scenario is when a Machine Learning model incorrectly predicts the identity of the employee from an image/video. These false positive scenarios can be detrimental to the credibility of Machine Learning solutions and can have negative impact on the experience and reputation of the concerned employee. Hence, such AI based solutions should be used to assist personnel in better monitoring of workplaces for adherence to guidelines and speed up their task, however, one should not remove the human from the loop completely.

# 8.
# Future Scope & Conclusion

Global regulations & laws regarding privacy and security of data collection and implementation of Facial AI systems vary around the globe. A single Facial AI system cannot be designed that can serve all and such systems need to be customized as per the regional data protection laws. The implementation of real-time face mask detector can impact some behavioral changes in employees. One can design A/B testing and focus groups studies, where one group will be monitored using Facial AI system, while the other group will not be. Such studies can help determine the extent of the impact such systems can have on human behavior in short to long term periods and thus eventually guide policy making at different levels from organizations to geographies. This type of study will reveal whether real-time monitoring affects in a positive or negative manner. The results will help in deciding the level of real-time employee monitoring which inspires optimal productivity.

**The solution can be further enhanced in the following ways:**

1. Model predictions can be saved in a database and used to build health monitoring dashboards capturing information such as number of employees following or violating Face Mask Guidelines, over a period of weeks and months, recognizing identity of individuals violating guidelines on regular basis and sending them appropriate reminders.
2. Integration with attendance tracking systems to ensure contactless data entry and with ERP system for automatic timesheet generation, thus saving time spent in data entry and record-keeping by every employee.
3. Better incident planning by CRO teams e.g., automate notifications to admin & emergency response teams if an employee collapses/faints in a secluded area, SOS calls, theft of office resources; thus an extension to such a system can act as a third-eye to CRO professionals.
4. Such a system can also help measure optimal utilization of space and help in better space planning of office premises, leading to savings in real estate cost.

5. When applied to other industries like retail, a facial recognition & object detection based AI monitoring system can help in activities like theft monitoring, inventory management, stock counting & planning.

In this paper, we have presented a system that uses Deep Learning to monitor healthcare guidelines, such as social distancing and face masking. We addressed the solution's technological design and how it may aid in the battle against COVID-19 and other such epidemics or pandemics. The paper also mentions implementing the same approach as a no-touch biometric system for attendance monitoring and automatic door opening, resulting in little human-to-surface interaction. We've also highlighted the application areas where a comparable solution may be applied, as well as the problems that come with it, such as data security and prediction accuracy.

# 9
# References


1. https://www.mohfw.gov.in/pdf/SOPonpreventivemeasurestocontainspreadofCOVID19inoffices.pdf
2. https://www.who.int/docs/default-source/coronaviruse/getting-workplace-ready-for-COVID-19.pdf
3. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide
4. https://arxiv.org/abs/1503.03832
5. https://www.kaggle.com/keras/resnet50
6. https://arxiv.org/abs/1506.01497
7. https://openai.com/blog/gpt-3-apps/
8. https://arxiv.org/abs/1810.04805
9. https://arxiv.org/pdf/1506.02640.pdf

# Authors

## Manish Shukla

*Assistant Manager – Data Science*

Manish Shukla is part of the Mphasis innovation and research group - Mphasis NEXT Labs. He holds an M. Tech. degree from Indian Institute of Technology, Kanpur in Industrial & Management Engineering. His technical skills include Machine Learning, Deep Learning, Operations Research, and Statistical Modelling. He has extensively worked in the field of Text Analytics, Natural Language Processing, Computer Vision, Email Order Processing, and Information retrieval from unstructured documents. He has a keen interest in advanced AI ML technologies and has been exploring different methods of Applied AI focused on NLP.

## Narayan Mishra

*Assistant Manager – Data Science*

Narayan Mishra is an Assistant Manager – Data Science at Mphasis. He completed his Master of technology from IIT Kanpur in Industrial & Management Engineering. He is proficient in Operations Research and Statistical Modelling. He is interested in various Machine Learning & Deep Learning areas like product recommendations, natural language processing, and computer vision. He has also been actively involved in building a production scale recommendation system for a large organization. He is exploring workflow management platforms for data engineering pipelines to programmatically author, schedule, and monitor workflows. He has a keen interest in Quantum Computing.

## Vibha Bhagchandani

*Associate Vice President, AI & Cognitive Solutions, Mphasis NEXT Labs*

Vibha Bhagchandani is part of Mphasis innovation and applied research group, NEXT Labs. Vibha is responsible for solution design and implementation of client-focused cognitive projects. She leads data science teams that deliver cognitive solutions built on cloud platforms and enable clients in their AI journey. She has been actively involved in the research, development, and delivery of product development initiatives focused on natural language processing (NLP) for some of Mphasis's premier clients. She is an enthusiastic salsa dancer and a foodie who loves trying world cuisine.

## About Mphasis

Mphasis (BSE: 526299; NSE: MPHASIS) applies next-generation technology to help enterprises transform businesses globally. Customer centricity is foundational to Mphasis and is reflected in the Mphasis' Front2Back™ Transformation approach. Front2Back™ uses the exponential power of cloud and cognitive to provide hyper-personalized ($C = X2C^{2}_{TM} = 1$) digital experience to clients and their end customers. Mphasis' Service Transformation approach helps 'shrink the core' through the application of digital technologies across legacy environments within an enterprise, enabling businesses to stay ahead in a changing world. Mphasis' core reference architectures and tools, speed and innovation with domain expertise and specialization are key to building strong relationships with marquee clients. To know more, please visit www.mphasis.com

**For more information, contact: marketinginfo.m@mphasis.com**

**USA**
460 Park Avenue South
Suite #1101
New York, NY 10016, USA
Tel.: +1 212 686 6655

**UK**
1 Ropemaker Street, London
EC2Y 9HT, United Kingdom
T : +44 020 7153 1327

**INDIA**
Bagmane World Technology Center
Marathahalli Ring Road
Doddanakundhi Village
Mahadevapura
Bangalore 560 048, India
Tel.: +91 80 3352 5000

www.mphasis.com

NR 28/11/21 US LETTER 9ASL7111